

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2005, 12:16:52 ; Search time 213 Seconds

(without alignments)
1876.966 Million cell updates/sec

Title: US-10-627-075-1

Sequence: 1 gagcttcaccagaccgacgc.....ccaaaaagagagagacgacg 1220

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.New.*
1: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44.6	3.7	96583	7 US-11-117-187-203	Sequence 203, App
2	42.4	3.5	1711	6 US-10-750-185-25798	Sequence 25798, A
3	41.6	3.4	1625	6 US-10-750-185-55754	Sequence 55754, A
4	40	3.3	1593	6 US-10-750-185-54114	Sequence 54114, A
5	39.4	3.2	496	6 US-10-131-826A-533	Sequence 533, App
6	39.2	3.2	156250	7 US-11-121-086-86	Sequence 86, App1
7	38.8	3.2	101786	7 US-11-117-187-199	Sequence 199, App1
8	38.6	3.2	159146	7 US-11-121-086-49	Sequence 49, App1
9	37.8	3.2	611587	7 US-11-117-187-209	Sequence 209, App
10	37.8	3.1	789	6 US-10-750-185-54249	Sequence 49249, A
11	37.8	3.1	988	6 US-10-750-185-61642	Sequence 61642, A
12	37.8	3.1	3631	6 US-10-750-185-61642	Sequence 61642, A
13	37.8	3.1	1082144	7 US-11-117-187-211	Sequence 211, App
14	37.8	3.1	1082144	7 US-11-117-187-211	Sequence 211, App
15	37.4	3.1	1450	6 US-10-750-185-52397	Sequence 52397, A
16	37.4	3.1	246960	7 US-11-121-086-8	Sequence 8, App1
17	37.4	3.0	2479	6 US-10-955-054A-50	Sequence 50, App1
18	36.8	3.0	1289	6 US-10-750-185-40390	Sequence 40390, A
19	36.8	3.0	2420	7 US-11-186-284-102	Sequence 102, App
20	36.2	3.0	1595	6 US-10-750-185-47422	Sequence 47422, A
21	36.2	3.0	1967	6 US-10-750-185-27185	Sequence 27185, A
22	36.2	3.0	3926	6 US-10-793-626-4300	Sequence 4300, App
23	36.2	3.0	3926	6 US-10-793-626-4300	Sequence 4300, App

24	36.2	3.0	4189	6	US-10-793-626-3898	Sequence 3898, App
25	36.2	3.0	4339	6	US-10-909-125-801	Sequence 801, App
26	36	3.0	2356	6	US-10-821-234-815	Sequence 815, App
27	36	3.0	161874	7	US-11-121-086-75	Sequence 75, App1
28	35.8	2.9	1073	7	US-11-102-240-95	Sequence 95, App1
29	35.8	2.9	184000	7	US-11-121-086-37	Sequence 37, App1
30	35.8	2.9	189252	7	US-11-121-086-54	Sequence 54, App1
31	35.8	2.9	189993	7	US-11-121-086-78	Sequence 78, App1
32	35.6	2.9	794	6	US-10-750-185-62887	Sequence 62887, A
33	35.6	2.9	1619	6	US-10-750-185-63154	Sequence 63154, A
34	35.6	2.9	3657	6	US-10-750-185-43919	Sequence 43919, A
35	35.6	2.9	6450	7	US-11-091-668-3	Sequence 3, App1
36	35.4	2.9	958	7	US-11-179-411-5	Sequence 5, App1
37	35.4	2.9	1608	6	US-10-750-185-24686	Sequence 24686, A
38	35.4	2.9	164527	7	US-11-121-086-71	Sequence 71, App1
39	35.2	2.9	742	6	US-10-750-185-62008	Sequence 62008, A
40	35.2	2.9	1720	6	US-10-750-185-43422	Sequence 43422, A
41	35.2	2.9	21777	7	US-11-027-964-2	Sequence 2, App1
42	35.2	2.9	43948	6	US-10-949-720-393	Sequence 393, App
43	35	2.9	959	6	US-10-750-185-60341	Sequence 60341, App
44	35	2.9	1267	6	US-10-750-185-54507	Sequence 54507, A
45	35	2.9	1468	6	US-10-750-185-57297	Sequence 57297, A

ALIGNMENTS

```
RESULT 1
US-11-117-187-203
; Sequence 203, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUS, DAPHNE
; APPLICANT: COPELHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 203
; LENGTH: 96583
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-203

Query Match      3.7%; Score 44.6; DB 7; Length 96583;
Best Local Similarity 46.8%; Pred. No. 0.17;
Matches 140; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

263 GTTCTCTCGACATCTTCACTCAATATCTATGTCTTAAACAGCAATTAATCAAG 322
59989 GTGTGATTTTCACTTTTGAGCATCTATTTGTCTATCTTATGATCATATATATGT 60048
322 CCAACCACTTAAGTTTATTTATGATATTTTGTATGACACAGACATGTAGTG 382
60049 CAAATTAATCTTAAGCACTCTTCAATTAAGAGTCAATTCAGCAATTAAGGATG 60108
383 CTGAGAAACGTATGTGTATCTTATTAATTAATCAAACTGGGTTAAATTAATAA 442
60109 AATTCAAAAGTTCGATTCAGCAATGATTAATGCAAAATGCAATTAAGCTAAGTAA 60168
443 AATGTGATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 502
60169 AATGATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 60228
503 AACCAATTAACAGGCTGATCTTGAAGCTTTATTTATTAATTAATTAATAAATAA 561
60229 CAAGATATTAAGAGCTCAGGCTTGAAGCTTTCTCAGAGATTAATAATGATGATATAG 60287
```

```

RESULT 2
US-10-750-185-25798/c
; Sequence 25798, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25798
; LENGTH: 1711
; TYPE: DNA
; ORGANISM: Bovine 19866881051295
US-10-750-185-25798

Query Match
Best Local Similarity 46.3%; Score 42.4; DB 6; Length 1711;
Matches 139; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 260 TTGTTTCCCTGACATCTTCACCTCAATATCTTTGTTACTTAACAGCAATTAATC 319
DB 1063 TTCTCTACCTCAATATACATGATTGAGAGAGGAGATVCAATTAATAATTTTAAAA 1004
QY 320 AGGCCAACCACTTAAAGTTTATTTGATATTTGTTAGGACAGACATGTGA 379
DB 1003 ATATATTTCCCTTTTCAATTAATTAATCTGCAATATATTAATGCTCAATTTGA 944
QY 380 GTGCTGAGAAACTGATGTTGTAATTAATTAATCAAACTGGGTTAAATAA 439
DB 943 CTCAGGTGAAAGCATGTGAATGATATCTCCAGAAAATCATGATTTTCAGAGTAA 884
QY 440 AAAAAATGCACTAACTTAAAAAAAACCAATACCAAGACCTTACTCCCTTGG 499
DB 883 AATAAATTCATGATGGGATATGATTAACAAGACAGTGGGAAATATGCAATTTGG 824
QY 500 AAAAGCACTTACAAGGCTGATCTTAACCTTTATTTTACATTAAGAAAATTAACA 559
DB 823 AAGAAAGTCTGTGAAGAAAGAGATTAATTTTAAAAACATTAACATGAGAAAAAGCA 764

RESULT 3
US-10-750-185-55754/c
; Sequence 55754, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 55754
; LENGTH: 1625
; TYPE: DNA
; ORGANISM: Bovine 19866881253827
US-10-750-185-55754

Query Match
Best Local Similarity 46.1%; Score 41.6; DB 6; Length 1625;
Matches 140; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 182 GGCATTAATTTAATCAGAAACATGCCACAGATGAGCCCGAGAAACCTGTAAAG 241
DB 452 GAGAAATTAATTTAAGCTGTGATTTCTACATAGGACGAGAGTGTCTCCAGAAAGAAAG 393
QY 242 TGAAGAAAAGGTATGTTGTTTCTCTGACATCTTCACTCAATATCTATTGTTA 301
DB 392 GAAATTAAGATTAATCTGATTTCTGATGATCCAAACATATTAATTAATTTGTTA 333
QY 302 CTTAAACAGCAATTAATCAGCCCAACCACTTAAGTTTATTTGATATGATTTGAT 361
DB 332 GAGAAATTCAGAAATGAAATGAGCAAAAGTAAATTAATTTCAAACTCAGATTAATTTGA 273
QY 362 TAAAGCAGACATGTGAGTCTGAGAAACTGATGTTGTTAATTAATTAATATC 421
DB 272 GAAATCATTAAAGTTGATTAAGAAAAGAAATGATCATAGTGAATCTTGACACAGCTG 213
QY 422 AAATGCGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481
DB 212 CAAAAAATTTGCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 153
QY 482 GACT 485
DB 152 AACT 149

RESULT 4
US-10-750-185-54114/c
; Sequence 54114, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54114
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Bovine 19866881235836
US-10-750-185-54114

Query Match
Best Local Similarity 58.3%; Score 40; DB 6; Length 1593;
Matches 70; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 372 ACATGTAGTGTGAGAAACATGATGTGTAATTAATTAATTAATTAATTAATTAAT 431
DB 1574 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1515
QY 432 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 491
DB 1514 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1455

```

```

Db      437 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 472

RESULT 6
US-11-121-086-86
; Sequence 86, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
APPLICANT: POULSEN, TIM S.
APPLICANT: NIELSEN, KIRSTEN V.
TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
FILE REFERENCE: 09138.6000-00000
CURRENT APPLICATION NUMBER: US/11/121,086
CURRENT FILING DATE: 2005-05-04
PRIORITY APPLICATION NUMBER: 60/567,570
PRIOR FILING DATE: 2004-05-04
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.3
SEQ ID NO 86
LENGTH: 156250
TYPE: DNA
ORGANISM: Homo sapiens
US-11-121-086-86

Query Match          3.2%; Score 39.2; DB 7; Length 156250;
Beet Local Similarity 47.5%; Pred. No. 6.7;
Matches 116; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY      389 AAACGTGATGTTGGTAAGTCTTGAATTAAATAATATCAAACTGGGTTAAAATAAAAAAATGT 448
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      4699 AAAATTACTTGTGTAAGTATATCTTACACAAGCCCTTAGTATCTCATAGAAATAT 4758
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      449 GCATTAAGTAAAAAACCATAATACCAAGACTTTACCTCCCTTGGAAGAAGCACA 508
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      4759 ATAAATATGTAAGATGATGAAATATTAATTAATTAATTAATTCCTCGGGAATTAATAT 4818
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      509 TTACAAGGCGTAGCTTACCTTTATATTAATTAACAATTAAGAAATAAACCAAGGTC CGA 568
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      4819 TCTAAATTTTGGATATATTTCTTCTGTTTTTAAAAAATATTCATCAACAGACAT 4878
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      569 TATAGCGTATATTTATTTCTCTAAGAGACAACTTCCTATGCTTAAATTAATTAAGT 628
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      4879 TATATTTTAAAGAGTTTGATGACCTTCAAGAACTTTCTTACATGTAATGTAAGAT 4938
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      629 GATT 632
        |||
DB      4939 GTTT 4942
        |||

RESULT 7
US-11-117-187-199/c
; Sequence 199, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
APPLICANT: PREUSS, DAPHNE
APPLICANT: COPELANDER, GREGORY
TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
FILE REFERENCE: ARCD.309US
CURRENT APPLICATION NUMBER: US/11/117,187
CURRENT FILING DATE: 2005-04-28
PRIORITY APPLICATION NUMBER: US/09/531,120
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US/09/531,120
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/125,219
PRIOR FILING DATE: 1999-03-18
NUMBER OF SEQ ID NOS: 212
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 199
LENGTH: 101786
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-11-117-187-199
```

```

Query Match
Best Local Similarity 3.2%; Score 38.8; DB 7; Length 101786;
Matches 91; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 290 TATCTATGTTACTTAACAGACAAATTAATCAGGCAACCACTTAAGTTTATTTAT 349
DB 12875 TATCTAGTAAGTTTCAACATTCGATTAATTTATTTATCTCTCGCATTTTATGTTAT 12816
QY 350 AGTATTTTGTGTAAAGCAGACACATGTAAGTGTGAGAAAATGATGTTGTAACCTGA 409
DB 12815 TTTTGTGTTTATTAAGACCGTTGATAGCGATTTTATTTATGATGATCTATTA 12756
QY 410 TTTAATATATCAACTGGGTTAAATATAAATAAATGTCATTAATAAAAAA 467
DB 12755 TTTTATTAATGGAAGCTGAGCATGATTTGTAATTTTTCATTAATTTAAAAA 12698

RESULT 8
US-11-121-086-49
; Sequence 49, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: FOLSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121.086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; LENGTH: 159146
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-49

Query Match
Best Local Similarity 3.2%; Score 38.6; DB 7; Length 159146;
Matches 80; Conservative 0; Mismatches 49; Indels 2; Gaps 1;

QY 387 GAAACTGATGTTGTTGTTACTGATTTAATATATCAAACTGGTTAAATAAAAAAAT 446
DB 69655 GAAAAATCCAGATGCTCTGATTTTTCATGTCAGCAACTTTAAATAAAAAAT 69714
QY 447 GTCATTAATTAATAAAAAAACAATATACCAACAGACTT--TACTTCCCTTGAAAAAG 504
DB 69715 TCCCAAAATAGTGAAGAGAAAACAATCTCATTTGCACTTCTTAATGGCCTTGAGCCA 69774
QY 505 CACATTTCACAA 515
DB 69775 CTGATTTCACAA 69785

RESULT 9
US-11-117-187-209
; Sequence 209, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPELHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117.187
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531.120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 209

```

```

; LENGTH: 611587
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-209

Query Match
Best Local Similarity 3.2%; Score 38.6; DB 7; Length 611587;
Matches 113; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 290 TATCTATGTTACTTAACAGACAAATTAATCAGGCAACCACTTAAGTTTATTTGTA 349
DB 201593 TAAATATTAATATATGTTTATTAATTAATATATATATTAATTAATTAATTA 201652
QY 350 AGTATTTTGTGTAAAGCAGACACATGAGTGTGAGAAAATGATGTTGTAACCTGA 409
DB 201653 TATGTTTGTGTTAAGTGAAGCACTTAATTAATCTTAATAAGAAAGATTTAAGATT 201712
QY 410 TTTAATATATCAAACTGGTTAAATATAAATAAATGTCATTAATAAAAAAACC 469
DB 201713 ATCTTAATTAATTAATTAATTTTCAATTAATAATTAATTAATTAATTAATTA 201772
QY 470 AAATACCAACAAAGCTTACTTCCCTTGGAAGAACACATTTACAGAGGCTGATCTT 526
DB 201773 AAGTCAAAATCAAAATTTATTTAATTAACAAAAACGTTTGAAGAAATATATT 201829

RESULT 10
US-10-750-185-54249
; Sequence 54249, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750.185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54249
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-54249

Query Match
Best Local Similarity 3.1%; Score 37.8; DB 6; Length 789;
Matches 84; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 526 TACCTTATATTTAATCAATTAAGAAAATAAACCAAGTCCGATATAGCTGTATTTAT 585
DB 209 TACCAATATCATTAAGCAATGAATAAATATCATTAAGAAAACCTATAGATTTCTAT 268
QY 586 TCCATAAAGAAAGAACTTTCATCTATGCTTTTAAATTAAGTATTAAGTATTAAGTAT 645
DB 269 GACCAATAGCCCAACAAATTCACCTAGATTAATTAATTAAGTATTAAGTATTAAGTAT 328
QY 646 TGCAGCTTAGCTTACAAACGAGCAGACAGACAGACAGAGCC 686
DB 329 AGGAGACATATGCTTACACAAAATTAATAACGAGAAACTGCC 369

RESULT 11
US-10-750-185-49243
; Sequence 49243, Application US/10750185
; Publication No. US20050260603A1

```


	Query Match	3.1%	Score 37.8	DB 6	length 3631
	Best Local Similarity	48.4%	Pred. No. 2		
	Matches 105	Conservative 0	Mismatches 112	Indels 0	Gaps 0
Qy	407	TGATTTAATATATCAAACTGGGTTAAAAAATGTCATACTTAAAAAAA			466
Db	2102	TGGTGAGATTCATTTTGATATTTGGCAAACTAATATCAATCATGTAAAGTTTAAAAATAA			2167

RESULT 14
US-11-117-187-211/c
Sequence 211, Application US/11117187
Publication No. US2005026560A1
GENERAL INFORMATION:
APPLICANT: PREUSS, DAPHNE
APPLICANT: COPENHAVEN, GREGORY
TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
FILE REFERENCE: ARCD:309US
CURRENT APPLICATION NUMBER: US/11/117,187
CURRENT FILING DATE: 2005-04-28
PRIOR APPLICATION NUMBER: US/09/531,120

;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/125,219
;; PRIOR FILING DATE: 1999-03-18
;; NUMBER OF SEQ ID NOS: 212
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 211
;; LENGTH: 1082144
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-11-117-187-211

Query Match 3.1%; Score 37.6; DB 7; Length 1082144;
Best Local Similarity 50.3%; Pred. No. 44;
Matches 93; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 274 ACATCTCTCACTCAAAATATCTATTGTTACTTAA CAGACAAATTAATCAGGCCAAACCACTT 333
DB 127984 ATATGATTTCTTTCACATGATTAATCACTCACTCACTTAAATTAATTCACAAAATTT 127925
QY 334 TAAGTTTATTTGATAGTATTTTGTGTTAAGGCACAGACATGTGAGTGTGAAAACT 393
DB 127924 AACAACTATGCTTAATTAATTAATTAATCAAACTGGTTAAATTAATTAATTAATGTCATA 127865
QY 394 GATGTTGGTAACCTGATTAATTAATTAATTAATCAAACTGGTTAAATTAATTAATTAATGTCATA 453
DB 127864 GAAGACAGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 127805
QY 454 ACTTA 458
DB 127804 AATTA 127800

RESULT 15

US-11-117-187-211
;; Sequence 211, Application US/11117187
;; Publication No. US2005026560A1
;; GENERAL INFORMATION:
;; APPLICANT: PREUSS, DAPHNE
;; APPLICANT: COPENHAVEN, GREGORY
;; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
;; FILE REFERENCE: ARCD:309US
;; CURRENT APPLICATION NUMBER: US/11/117,187
;; CURRENT FILING DATE: 2005-04-28
;; PRIOR APPLICATION NUMBER: US/09/531,120
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/125,219
;; NUMBER OF SEQ ID NOS: 212
;; SOFTWARE: Patentin Ver. 2.1
;; SEQ ID NO 211
;; LENGTH: 1082144
;; TYPE: DNA
;; ORGANISM: Arabidopsis thaliana
US-11-117-187-211

Query Match 3.1%; Score 37.6; DB 7; Length 1082144;
Best Local Similarity 48.6%; Pred. No. 50;
Matches 103; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 414 ATAAATATCAAACTGGGTTAAATAAAAAAATGTCATTAATAAAAAAACAAT 473
DB 556760 AAAATATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 556819
QY 474 ACCAACAAGACTTCTCCCTTGGAAAAGCATTTCAGAGGCTGATCTAGCCCTT 533
DB 556820 GATATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 556879
QY 534 ATATTTACAATTAAGAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 593
DB 556880 ATATGTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 556939
QY 594 GAACGAAACTTTCATGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTC 625

DB 556940 AAACATAAATTAATAAAAAATGAGAGAAAAGAA 556971

Search completed: December 12, 2005, 15:45:07
Job time : 249 secs

Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landere, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, C., Spencer, P., Spencer, B., Stange-Thoman, N., Stojanovic, N., Stubb, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (07-MAR-2004) Whitehead Institute/MIT Center for Genome Research 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 228046)

REFERENCE
AUTHORS
Britten, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, M., Anderson, S., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collumore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landere, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, C., Smith, C., Spencer, B., Stange-Thoman, N., Stojanovic, N., Stubb, M., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (06-APR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 6, 2004 this sequence version replaced gi:42794163.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@brad.mit.edu
Project Information
Center project name: L23266
Center clone name: 50_L_10

FEATURES
source
Location/Qualifiers
1..228046
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="8"
/map="8"
/clone_lib="RP23-50L10"
/note="wgs_end extension
1..408
clone_end:SP6"
174..220
/rpt_family="(TG)n"

misc_feature
409..414
/note="clone boundary
clone_end:SP6
site:BCRI"
complement(1173..1303)
/rpt_family="B1_MM"
1757..1805
/rpt_family="GA-rich"
1822..2158
/rpt_family="ORR1A2"
2162..2243
/rpt_family="(AATAG)n"
complement(3076..3723)
/rpt_family="RMR19B"
3762..3809
/rpt_family="(CNGA)n"
complement(3816..3956)
/rpt_family="B1_MM"
complement(4454..4725)
/rpt_family="MLT1C"
complement(4769..4924)
/rpt_family="B3"
4946..5135
/rpt_family="MTB"
5151..5503
/rpt_family="MTD"
complement(5904..5997)
/rpt_family="B4A"
7034..7172
/rpt_family="RMR15"
8071..8127
/rpt_family="(TG)n"
8228..8339
/rpt_family="PB1D9"
8351..8399
/rpt_family="A-rich"
9201..9283
/rpt_family="(GGGTG)n"
complement(9360..9438)
/rpt_family="MER34"
complement(10018..10162)
/rpt_family="MT-Int"
10172..10326
/rpt_family="B2_Mm2"
10329..10374
/rpt_family="(CAA)n"
complement(10383..11039)
/rpt_family="WT-Int"
complement(11040..11231)
/rpt_family="B2_Mm1"
complement(11232..11376)
/rpt_family="WT-Int"
complement(11377..11569)
/rpt_family="B2_Mm1"
complement(11570..11729)
/rpt_family="WT-Int"
complement(11732..12136)
/rpt_family="WTC"
complement(12182..12333)
/rpt_family="URR1A"
12495..12642
/rpt_family="B1_MM"
12654..12699
/rpt_family="(GAAA)n"
complement(13350..13517)
/rpt_family="MLT1A"
13789..13831
/rpt_family="(T)n"
14745..14779
/rpt_family="MIR"
complement(14962..15112)
/rpt_family="ORR1D"
15389..15509

repeat_region /rpt_family="TA)n"
15535 .15602
/rpt_family="(CATATA)n"
repeat_region 15604 .15705
/rpt_family="PB1D7"
repeat_region 15706 .15772
/rpt_family="ID3"
repeat_region complement(15813 .16133)
/rpt_family="MTB"
repeat_region complement(16623 .16776)
/rpt_family="Charlies"
repeat_region 17613 .17728
/rpt_family="PB1D10"
repeat_region 18154 .18320
/rpt_family="B3A"

Query Match 99.9%; Score 1218.4; DB 9; Length 228046;
Best Local Similarity 99.9%; Pred. No. 1.8e-270;
Matches 1219; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGCTTCCAGTACCGATCTCTGCAGCATTAATTCTAGTCATGAAAGGGGTGTGACTC 60
DB 165469 GAGTCTCCAGTACCGATCTCTGCAGCATTAATTCTAGTCATGAAAGGGGTGTGACTC 165528
QY 61 TGGACGAGAGAGCCAGAGCTGGAAGCTTTTAAAGTCAGTGAAGCCCAAGCTTTT 120
DB 165529 TGGACGAGAGAGCCAGAGCTGGAAGCTTTTAAAGTCAGTGAAGCCCAAGCTTTT 165588
QY 121 ATGGCAGAGGGGCTTTGAACTTAATTAAGGGAAGTATTTGCTGAGCCCACTGAC 180
DB 165589 ATGGCAGAGGGGCTTTGAACTTAATTAAGGGAAGTATTTGCTGAGCCCACTGAC 165648
QY 181 TGGGACTAATTTATCAGGAACATGCCAGAGTATGAGCCCGAGAAACCTGATACA 240
DB 165649 TGGGACTAATTTATCAGGAACATGCCAGAGTATGAGCCCGAGAAACCTGATACA 165708
QY 241 GGAAGGAAGAGGTATGTTTCTCTCGACATCTTCACTCAATATCTATTTT 300
DB 165709 GGAAGGAAGAGGTATGTTTCTCTCGACATCTTCACTCAATATCTATTTT 165768
QY 301 ACTTAACAGACATTAATCAGGCAACCACTTAATTTTATAGTATTTT 360
DB 165769 ACTTAACAGACATTAATCAGGCAACCACTTAATTTTATAGTATTTT 165828
QY 361 TTAAGGCACAGATGTGAGTGTGAGAAAGTATGTTGATTAATTAAT 420
DB 165829 TTAAGGCACAGATGTGAGTGTGAGAAAGTATGTTGATTAATTAAT 165888
QY 421 CAATCTGGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 165889 CAATCTGGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 165948
QY 481 AGACTTTACTCCCTTTGAAAAAGACATTTTACAGGGCTGATCTTAGCTTTATATTA 540
DB 165949 AGACTTTACTCCCTTTGAAAAAGACATTTTACAGGGCTGATCTTTATATTA 166008
QY 541 CAATAAAGAAATTAACCAAGTCCCGATATAGCTGTAATTTTCTTAAAGACAGA 600
DB 166009 CAATAAAGAAATTAACCAAGTCCCGATATAGCTGTAATTTTCTTAAAGACAGA 166068
QY 601 AACTTTCACTAGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 660
DB 166069 AACTTTCACTAGCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 166128
QY 661 CAAACGAG 720
DB 166129 CAAACGAG 166188
QY 721 CTCTGACATGCCAGAGGGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 166189 CTCTGACATGCCAGAGGGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 166248
QY 781 CAGTCTTTGGGGGAGTAAGCATAGCTGATTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 840

DB 166249 CAGTCTTTGGGGGAGTAAGCATAGCTGATTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 166308
QY 841 CCTACAGAGAGATAGTGGGTGAGCCAGGGGGGAGAGCGGCTGCTGACATGTCTGGCTG 900
DB 166309 CCTACAGAGAGATAGTGGGTGAGCCAGGGGGGAGAGCGGCTGCTGACATGTCTGGCTG 166368
QY 901 CTCTTATCACTTATATTAATTAAGGAGAGAAAGTATTAATTCGATCTGACATCTGTAG 960
DB 166369 CTCTTATCACTTATATTAATTAAGGAGAGAAAGTATTAATTCGATCTGACATCTGTAG 166428
QY 961 ACTCAGGGGAG 1020
DB 166429 ACTCAGGGGAG 166488
QY 1021 TCTCAGAGAGCTCCGAGTGTCCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 166489 TCTCAGAGAGCTCCGAGTGTCCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166548
QY 1081 GAGCTGTCTGGGAG 1140
DB 166549 GAGCTGTCTGGGAG 166608
QY 1141 GTGCTTACAGAGAGCTGAGAGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 166609 GTGCTTACAGAGAGCTGAGAGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 166668
QY 1201 CCAAAAAGAGAGAGAGATG 1220
DB 166669 CCAAAAAGAGAGAGAGATG 166688

RESULT 2
AC114046 228774 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-18728, WORKING DRAFT SEQUENCE.
DEFINITION AC114046
AC114046 GI:30581607
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 228774)
Muzny,D,Marie, Metzker,M,Le,., Adamson,S,., Adams,C,., Alder,J,., Allen,C,., Allen,H,., Alsbrooks,S,., Amin,A,., Anguiano,D,., Anyalebechi,V,., Ayogagi,A,., Ayodeji,M,., Baca,E,., Baden,H,., Baldwin,D,., Bandaramaika,D,., Barber,M,., Barnstead,M,., Benahmed,F,., Biswal,K,., Blair,J,., Blankenburg,K,., Blyth,P,., Brown,M,., Bryant,N,., Buhay,C,., Burch,P,., Butrell,K,., Calderon,E,., Cardenas,V,., Carter,K,., Cavazos,I,., Ceasar,H,., Center,A,., Chacko,J,., Chavez,D,., Chen,G,., Chen,R,., Chen,Y,., Chen,Z,., Chu,J,., Cleveland,C,., Cockrell,R,., Cox,C,., Coyle,M,., Cree,A,., D'Souza,L,., Davila,M,L,., Davis,C,., Davy-Carroll,L,., De Anda,C,., Dederich,D,., Delgado,O,., Denison,S,., Deramo,C,., Ding,Y,., Dimh,H,., Divya,K,., Draper,H,., Dugan-Rocha,S,., Dunn,A,., Durbin,K,., Duval,B,., Eaves,K,., Egan,A,., Becotto,M,., Eugene,C,., Evans,C,., Falls,T,., Fan,G,., Fernandez,S,., Finley,M,., Flagg,N,., Forbes,L,., Foster,M,., Foster,P,., Fraser,C,., Gabisi,A,., Gante,R,., Garcia,A,., Garner,T,., Garza,M,., Gebregeorgis,B,., Geet,K,., Gill,R,., Grady,M,., Guerra,M,., Guevara,M,., Gunaratne,P,., Haaland,W,., Hamil,C,., Hamilton,C,., Hamilton,K,., Harvey,Y,., Havlak,P,., Hawes,A,., Henderson,N,., Hernandez,J,., Hernandez,R,., Hines,S,., Hladun,S,., L., Hodgson,A,., Hogues,M,., Hollins,B,., Howell,S,., Huijck,S,., Hume,J,., Idlerbird,D,., Jackson,A,., Jackson,L,., Jacob,L,., Jiang,H,., Johnson,B,., Johnson,R,., Jolyet,A,., Karpathy,S,., Kelly,S,., Kelly,S,., Khan,Z,., King,L,., Kovar,C,., Kowis,C,., Kraft,C,., Lebow,H,., Levan,J,., Lewis,L,., Li,Z,., Liu,J,., Liu,J,., Liu,W,., Liu,Y,., London,P,., Longacre,S,., Lopez,J,., Lorenshew,L,., Loulseged,H,., Lozado,R,., Lu,X,., Ma,J,., Maheshwari,M,., Mahindaratne,M,., Mahmoud,M,., Malloy,K,., Mangum,A,., Mangum,B,., Mapua,P,., Martin,K,., Martin,R,., Martinez,E,., Mahoney,S,., McLeod,M,., McNeill,T,., Meenen,B,.,

Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Mundasa M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nwaelelemeh O., Okwunonu G., Olarnpusagoon A., Pal S., Parks K., Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C., Plopper F., Polindexter A., Popovic D., Primus E., Pu L., Puzos M., Quirroz J., Rachlin E., Reeves K., Regier M.A., Reigh R., Rives C., Rodkey T., Rojas A., Rose M., Rose R., Ruiz S.J., Sanders M., Saverly G., Scherer S., Scott G., Shatsman S., Shen H., Shetty U., Shvartsbeyn A., Sisson I., Sitter C.D., Smajs D., Sneed A., Sodergren E., Song X.-Z., Sorelle R., Soosa J., Steinle M., Strong R., Sutton A., Svatek A., Tabors P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K., Valas R., Vera V., Villasana D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren R., Wei X., White F., Williams G., Willson R., Wleczky R., Wooden H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu F., Zhang J., Zhou J., Zhou X., Zhou S., Dunn D., von Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.

Direct Submission
Unpublished
2 (bases 1 to 228774)
Worley K.C.

Direct Submission
Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 228774)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 13, 2003 this sequence version replaced gi:22857247. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/ret/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKMK
Center clone name: CH230-187J8
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 215732 bases at least Q40
Consensus quality: 217275 bases at least Q30
Consensus quality: 218412 bases at least Q20
Estimated insert size: 225110; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Gendbmk_drafc_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced

* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 228774: contig of 228774 bp in length.
Location/Qualifiers
source 1. 228774
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-187J8"
misc_feature 1. 1046
/note="wgs contig"
misc_feature 227275..228774
/note="wgs_contig"

ORIGIN
Query Match 65.0%; Score 793.4; DB 14; Length 228774;
Best Local Similarity 85.0%; Pred No. 2,1e-117;
Matches 1079; Conservative 0; Mismatches 111; Indels 79; Gaps 14;

2 AGCTTCCAGTACCGATCTCTGAGCATTTAATTCTAGTCATGAAAGGGTGTGACTCT 61
|||||
82805 AGCTTGCAACAGTACCAATCTCTGAGCATTTAATTCTAGTCATGAAAGGGTGTGACTCT -- 82862
62 GGACGAGAGAGCCACAGAGCTGGAAGTCT--TTAGAGTCAATGACCCCGAGCTTTT 120
|||||
82863 ---GGAACGAGAGCCACAGAGCTGGAAGTCTTTAGAGTCAATGACCCCGAGCTTTT 82919
121 ATGGCAGAGGGCTTTGGAATTTAATAAGGGGAATGATTTGCTGAGCCGACTGAC 180
82920 ATGGCCAAAGGGTTTGAATTTAATAAGGGGAATGATTTGCTGAGCTGACTGATGAC 82979
181 TGGGCTAATTTATATGAGAAATGCCACAGAGTATGAG--CCGAGGAAACCTGATATC 239
82980 TGAGCTAATTTAATACATTAATGCCACAGAAATGATGAGCCGAGGAAACCTGATATC 83039
240 AGTGAAGGAAAGGATGATTTGTTGTTCTCTGACATATCTCACTCAATATATCTATG 299
83040 AGCGAAGGAAAGGATATG---TGTTCTCTGATATCTCACTCAATATATCTATG 83095
300 TACTTAACAGCAATTAATGAGCCAAACCACTTAAGTTTATTTATGATATGTTTGT 359
83096 TACTTAACAGCAATTAATGAGCCAAACCACTTAAGTTTATTTATGATATGTTTGT 83155
360 G-----TTAAGGACAGCATGTGAGTCTGAGAAACATGATGTTGTTGTTGTTGA 409
83156 GTATTCTGTGTTAAAGTTAAACATGATGATGCTGAG--AAGCTGACCTGATTAATCTGA 83213
410 TTTAATATATCAAACTGGGTTAAAT---AAAAAAATGTGATTAATTAATAAAAAA 466
83214 TTTAGTAAATACCAATTTGGTTAAATTTAAATAAAAAAGTGCATTAATGATGATGA 83272
467 ACCAAATACCAAAAGACTTACTTCCCTTGAAAGCACTTTACAAAGGCTGATCTT 526
83273 ---AAAAACAGCAAACTTACTTCCCTTGAAAGCACTTTAAAGGCGGATCTT 83329
527 AGCTTATATTTATCAATTAAGAAATTAACCAAGTCCGATTTAGCTGTAATTTATT 586
83330 ACTCTTATATTTAATTAAGAAATTAACCA--GTCCTAATTTACTGTAAGTTTATT 83387
587 CCT-----AAAAAGACAGAACTTTGACTATGTTTAAATTT 623
83388 CTTAAACAAACAAACAAACAAACAAACAAACAAACAAACCTTCCATGCTTTAAATTT 83447
624 AAAGTATTTACTCTGATCTGCAAGTTAGCTTACCAACAGAGCAGACAGACAGACAG 683
83448 ACACTGATTTAACCAAGTATCTGCAAGCTCAGCTCAATATGAGCAGACAGACAGACAG 83507
684 GCCCAGCTCTCTCTGGAATTAATTAAGGTTGTTGCTGCAATGCCAGAGGCTTTG 743
83508 GCCCAGCTCTCTCTGGAATTAATTAAGGTTGTTGCTGCAATGCCAGAGGCTTTG 83567
744 TGAGTGTGTGTGTTCCAGAAAGCTTCTGACATACAGATCTTTGGGACAGTAAGCAC 803
|||||

Db	Accession	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
Db	83568	TGGCTGGTCTGTCTGTTCCCGAAGGCTTCTGTGAAGTGCACAGTCTTTGGGGCAGTATGAC	83662						
Qy	804	TATGCTGATATTTTCTGTTGCTGGCTAGTGAACCCCTTACAGAAATAGTGGGTGAG	863						
Db	83628	TATGCTGATATTTTCTGTTGCTGGCTAGTGAACCCCTTACAGAAATAGTGGGTGAG	83687						
Qy	864	CCAGGGGGCGGAGCGGCTGGCTGACATGTCTGGCTGCTTTATGCAATTATATTAAG	923						
Db	83668	CCAGGGGGCGGAGCGGCTGGCTGACATGTCTGGCTGCTTTATGCAATTATATTAAG	83747						
Qy	924	GGAAGAAAGTATTTGATTCGGATATCTGACATCTGTAGACTCAGGGGGAACAAAGAGTC	983						
Db	83748	GGAAGAAAGTATTTGATTTGGATATCTGACATCTGTAGACTCAGGGGGAACAAAGAGTC	83807						
Qy	984	CGTGCAGACCTCTGGAATGACAGGCGTCTCTCTCTCTCAGGACAGCTCGAGTGTGC	1043						
Db	83808	CGTGCAGACCTCTGGAATGACAGGCGTCTCTCTCTCTCAGGACAGCTCGAGTGTGC	83867						
Qy	1044	CGGGGAGAGAAAGAAAGAAAGAGCAGGCGCTGGGAAAGACCTGTGGGGAGCGAGAG	1103						
Db	83868	CGG-----GAGAAAGAGAGAGCCTGGGAAATCAGCTGCTTACG-----GACCGGGAGAG	83917						
Qy	1104	GCTCTCACTGATGAGACTTATTCACACGGGCAACAGCCCTGTGCTCTTACAGACAGCTAGAG	1163						
Db	83918	GCTCTCACTGATGAGACTTATTCACACGGGCAACAGCCCTGTGCTCTTACAGACAGCTAGAG	83975						
Qy	1164	CTCAGAGAC-----GCAAGTTTGTGAACTTCACAGTTTAAACCAACCAAGAGAG	1211						
Db	83976	CTCAGAGACTCTTATGTGCTGCTGCAAGTTGTGCTGAACCTCAGTTTAAACCAACCAAGAGAG	84035						
Qy	1212	GAGAGAAATG 1220							
Db	84036	AAGAAATGTG 84044							
RESULT 3	AC132960								
LOCUS	AC132960								
DEFINITION	AC132960 326593 bp DNA linear HTG 19-NOV-2002								
ACCESSION	AC132960								
KEYWORDS	HTGS, HTGS_PHASE1, HTGS_DRAFT, HTGS_ENRICHED.								
SOURCE	Rattus norvegicus (Norway rat)								
ORGANISM	Rattus norvegicus								
REFERENCE	Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi, Mammalia, Eutheria, Euarchontoglires, Glires, Rodentia; Sciurognathi, Murioidea, Muridae, Murinae, Rattus.								
AUTHORS	1 (bases 1 to 326593)								
	Mullen, D., Marle, M., Metzger, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooke, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhammed, F., Bismato, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyte, M., Cree, A., D'Souza, L., Davila, M., L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Diaper, H., Dugan-Kocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Gante, R., Garcia, A., Garner, T., Garza, M., Geleorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guerra, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulvik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwie, C., Kraft, C.L., Lebow, H., Lervan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,								

Mareshwari, L., Louisaged, H., Lozada, R. J., Lu, X., Ma, J.,
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
 Mahoney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
 Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Muidasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokeme, O., Okunnu, G., Olarinmason, A., Pal, S., Parks, K.,
 Patel, N. S., Paul, H., Perez, A., Perez, L., Plankoch, C.,
 Plapp, F., Poudexter, A., Popovic, D., Primus, E., Pu, L.,
 Puzo, M., Quirco, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
 Rell, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
 Sanders, W., Savary, G., Scherer, S., Scott, C., Shatsman, S., Shen, H.,
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitzer, C. D., Smales, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Swatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villaena, D., Waldron, L., Walker, B., Wang, J.,
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
 Weissrock, G. and Gibbs, R. A.
 Rat Genome Sequencing Consortium.
 Direct Submission
 Unpublished
 2 (bases 1 to 326993)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (05-SEP-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 326993)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23322544.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KCDZ
 Center clone name: CH230-16203
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 276174 bases at least Q40
 Consensus quality: 282235 bases at least Q30
 Consensus quality: 286810 bases at least Q20
 Estimated insert size: 281736, sum-of-contigs estimation
 Quality coverage: 5x in Q20 bases, sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 16 contigs. The true order of the pieces


```
/clone="RP11-16G12"
/clone_lib="RP11-1 Human Male BAC"
repeat_region
/rpc_family="AluSg/x"
1..221
complement(1277..1428)
repeat_region
/rpc_family="L1ME1"
1732..1823
/rpc_family="(CA)n"
2184..2238
/notes="PCR product sequence only"
complement(2649..2774)
repeat_region
/rpc_family="L2"
2963..3133
/rpc_family="(TG)n"
3167..3208
/rpc_family="(TG)n"
3581..3587
/notes="<30 qual SNGL region"
complement(6732..6840)
repeat_region
/rpc_family="L2"
complement(8405..8540)
/rpc_family="MIR3"
8922..9027
/rpc_family="L2"
9061..9174
/rpc_family="L2"
9390..9595
/rpc_family="L2"
complement(9592..9699)
/rpc_family="MIR"
1109T..11333
/rpc_family="AluSx"
11359..11649
/rpc_family="AluSp"
11765..12100
/rpc_family="AluUb"
12325..12593
/rpc_family="AluUb"
13535..13813
/rpc_family="AluSp"
13817..13850
/rpc_family="(CAA)n"
13857..14036
/rpc_family="MTR"
17230..17345
/rpc_family="L1ME1"
17528..17551
/rpc_family="AT-rich"
1847T..18757
/rpc_family="AluUo"
21235..21353
/rpc_family="L2"
21420..21484
/rpc_family="L1PA9"
22294T..22456
/rpc_family="MTR"
complement(22452..22513)
/rpc_family="MIR"
22668..23024
/rpc_family="MLT1A1"
23390..23520
/rpc_family="MER85"
23535..23739
/rpc_family="MER2"
23908..24013
/rpc_family="MIR"
24043..24073
/rpc_family="(TTTG)n"
complement(24079..24356)
/rpc_family="AluSg"
24877..24952
/rpc_family="L1TR67"
27866..28203
repeat_region
```

```
repeat_region
/rpc_family="AluSx"
28205..28272
/rpc_family="GA-rich"
28663..28682
/rpc_family="(T)n"
complement(28688..29140)
repeat_region
/rpc_family="MLT1C"
2955T..29760
/rpc_family="L2"
complement(29782..30036)
/rpc_family="L1ME4"
30200..30591
/rpc_family="MLT1B"
complement(30615..30910)
/rpc_family="L1ME4"
3138T..31634
/rpc_family="L2"
31679..31699
/rpc_family="AT-rich"
31939..31990
/rpc_family="L2"
33132..33161
repeat_region

Query Match 20 0% Score 243.4; DB 8; Length 191377;
Best Local Similarity 64.2% Pred. No.1.6e-45;
Matches 554; Conservative 0; Mismatches 271; Indels 38; Gaps 11;

QY 206 CCACAGAGTGTGAGCCCGAGAAACCCGTGATACAGTGAAGAAAGGTATGTTGTT 265
DB 40710 CGACGAGCGAGACGTCGTCACAAAAAAGGTTGTTGTTGTTGTTGTTGTTGTT 40769
QY 266 TCCTTCGACATCTTCTCAATATCTATTTGTTACTTAAACAGCAATTATCAGGCCA 325
DB 40770 TCTACACAGGACCTTAATCTTAATTTCTCATGTTAATGTAATGACCTTAATCAGTCA 40829
QY 326 AACCACTTAAGTTTATTTGATAGTATTTGTTGTTAAGCAGACATGAGTGTGCTG 385
DB 40830 GTGTATTTTGAAGTGAATATCCAGTGAAGTGTGTTTAAAGA-ATAAGGTCAAGCTC 40888
QY 386 AGAAAACTGATGTTGGTGAATCTGATTATATAA-----TATCAACTGGGTTAAA 435
DB 40889 AG-AAAGTGAATGTTCTAAGCGTGAATATAACAGATCTTCTATCAAGCTGTTGGAAA 40947
QY 436 TAAAAA-----AAATGTGATATCTTAAAAAACCATAACCAACAGACTTTA 488
DB 40948 TGAATATCTTTTGTAACTGTGCTAGTGTGTCAGATCACTAAATCAATGAATCTATTT 41007
QY 489 CTTCCTTGGAAAAAGCATTTACAGAG--CTGATCTTAGCTTTATATTTACATA 545
DB 41008 TTTCCCTCTGAAGAAAGCAGAGTTTACAGAGCCAACTTCTTCTCTTATTTGTAATTA 41067
QY 546 AAGAAATTAACCAAGGTCCCGATATAGCTGAATTTATTCCTA-----AAGAAACA 598
DB 41068 ACAAATATTAACCGAAGTCTGACCTATTTGAATATTTTATTCCTAAGAAAAAACA 41127
QY 599 GAAATTTTCACTATGCTTTAAATTAAGTATTAACCTCAATACTCTGCAAGCTTACCC 658
DB 41128 GGAATTTTCACTATGCTTTCAATTAAGTATTAACCTCAATATTTTGGCCAGCTTAACA 41187
QY 659 TACAAAGAGCAGA--CAGACAAAGAGCCCAAGCT-ACCTCTGCAAAATATTAATTA---- 711
DB 41188 CGGCAAAAAATCAGTTTCAAGACAAAGAGATCAATGCTCTCTCTGAGAAATATCTTAATG 41247
QY 712 GGGTGTGCTCTGACATGCCAGGGGTCTTGTGGCTGTG- TGTTCCAGAAAGCTT 770
DB 41248 GGGTGTGCTCTGAGAAATGCCAGGGGTCTGTAACAATCGGTTTTTCCAGAGAGGTTT 41307
QY 771 CTGACATACAGATGCTTTGGGCGAGTAAAGCATATGCTGATTTTTCGTTGCCCTGG 830
DB 41308 CTGACGATGGGTCTGTGTGAGGGCGAGGATTTCTGATTTTTCGTTGCCCTGG 41367
QY 831 CTAGTACCCCTACAGAGAAATAGTGGTGAAGCAGG-GGGCGAGCGGCTGCTGAC 889
```


	Best Local Similarity	65.1%	Pred. No. 1.7e-44;	
	Matches 537;	Conservative 0;	Mismatches 250;	Indels 38; Gaps 11;
QY	244	AAAGAAAAGGTGATGTTGTTGTTCCCTGACATACCTCACTCACTCAATATCTATGTTACT	303	
Db	118402	AAAAAAGTTGTTATCTCTCTTCTACACAGACGCTTAATTTCTATTTCTATTTGTA	118344	
QY	304	TAAACAGCATTAATACGCGCAACCACTTAAAGTTTATTTATTTGTAATTTTGCTTA	363	
Db	118342	TGTAATGAACTTAATCAAGTACAGTATTTTATTTGAGAGTAATCAAGTGAAGTTTTTT	118283	
QY	364	AGGCACAGACATGTGAGTGTGAGAAAACATGATGTGTACTGATTTATATA-----	417	
Db	118282	AAGA-AATAAAGTCAGAGCTGAG-AACTGATGTTTCTTACCTGACTTAATTAACAGATC	118222	
QY	418	----TATCAACTGGGTTAAATTAATA-----AAATGTGCATTAATTAATAAAAAA	466	
Db	118224	TTTCTATCCAACTGTTTCGAAATGAATAATCTTTTGTAACTTGGCTGATGTTCAAGATC	118165	
QY	467	ACCAATATACCAACAGACTTACTTCCCTGGAAAAGCACTTACAAAGG---CTGAT	523	
Db	118164	AACATAATATCAATGATCTATTTTCCCTCTGAAAAGGACACAGTTTACAGAGCCAAAC	118105	
QY	524	CTTACGCTTTATATTTATACATAAAGAAAATAAACCAAGGTCCGATATAGCTGTAA	583	
Db	118104	TTCTTCCTCTTTATTTGTAATATAAATAATTAACGAAGTCCGACCTATTTGTATATT	118045	
QY	584	ATTCCCA-----AAGAACAAGAACTTTCACTATGCTTTAAATTTAAAGTATTACT	636	
Db	118044	TTTATTCCTTAAGAAAAGAAAACAGGAATTTGATTTGACTTCAACATTAAGTTATTACCT	117985	
QY	637	CAGATACCTGCAAGCTTAGCCTACAAACAGACAG--ACAGACACAGAGCCCA-GCTA	693	
Db	117984	CAGATATTTTGCAGCTTAGACACGGAAGAAATACAGTTTACAGACAAAGAGATCACTGCT	117925	
QY	694	CTCTCTAGAAA---TAAATTAAGGTGTGCTCTTGACATGCCACAGGGCTTTGTGCTG	749	
Db	117924	CTCTCTAGAAAATTAATTAATTTGGGGTGTGCTTAGAGAAATGCCAGGGGTCTGTAAACG	117865	
QY	750	GTCGTG--TTCCACAGAGGCTCTGCAGTACACAGTCCCTTGGGGAGTAAAGCACATATGC	808	
Db	117864	ATCGGTTTTTCCACAGAGGTTTTCTGCAGATAGGATCTGTGGAGGAGGACATTTCTGC	117805	
QY	809	TCTGATTTTTCTGTGTTCCCTGCTAGTGAACCCCTTACAGAAAGATAGTGGTGAGCCAG	868	
Db	117804	TCTGATTTTTCTGTGTTCCCTGCTAGTGAACCCCTTACAGAAAGATAGTGGTGAGCCAG	117745	
QY	869	-GGGGGAGGCGGCTGCTGCTGCAATGTCTGCTGCTTATATCAACTTATATTAAGGAA	927	
Db	117744	AGGGGAGGAGCCCACTACACATATCTGGGCTCTTATCAACTTATATTAAGGAA	117685	
QY	928	GGAAAGTATGATTCGATACGATACGACCTGTGACCTCAAGGGGAGAAACAAAGAGTCCGTG	987	
Db	117684	GGAAAGTATGATTCGATACGATACGATACGATACGATACGATACGATACGATACGATAC	117625	
QY	988	CAGACCTCTGAGTGAAGAGAGGGGCTCTCTTCTCTCAAGACAGC	1032	
Db	117624	GTGAAGCTGCTCTGTAAAGCTGACACAGCCCTCCCAAGTGAAGC	117580	
RESULT 6				
AX087869/c				
LOCUS	AX087869	240825 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0114550.			
ACCESSION	AX087869			
VERSION	AX087869.1	GI:13396862		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Hominidae; Homo.			
REFERENCE				
1				

AUTHORS Barry,C., Chumakov,I. and Blumenfeld,M.
TITLE Prostate cancer-related gene 3 (pg3) and diallelic markers thereof
JOURNAL Patent: WO 0114550-A 1 01-MAR-2001;
GENSET (FR)
FEATURES
source Location/Qualifiers
1..240825
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature
1..2000
/note="5,regulatory region"
primer_bind
1823..1840
/note="5-390.pu"
primer_bind
1980..1998
/note="5-390-177.mis"
misc_binding
1987..2011
/bound_moiety="5-390-177.probe"
variation
1999
/note="5-390-177 : polymorphic base G or C"
primer_bind
2000..2018
/note="5-390-177.mis complement"
exon
2001..2079
/note="exon A"
primer_bind
2108..2125
/note="5-390.rp complement"
primer_bind
4559..4577
/note="5-391.pu"
primer_bind
4582..4600
/note="5-391-43.mis"
misc_binding
4589..4613
/bound_moiety="5-391-43.probe"
variation
4601
/note="5-391-43 : polymorphic base A or G"
primer_bind
4602..4620
/note="5-391-43.mis complement"
exon
4627..4718
/note="exon B"
primer_bind
4891..4908
/note="5-391.rp complement"
primer_bind
10007..10025
/note="5-392.pu"
exon
10115..10213
/note="exon C"
primer_bind
10209..10227
/note="5-392-222.mis"
misc_binding
10216..10240
/bound_moiety="5-392-222.probe"
variation
10228
/note="5-392-222 : polymorphic base G or T"
primer_bind
10229..10247
/note="5-392-222.mis complement"
primer_bind
10267..10285
/note="5-392-280.mis"
misc_binding
10274..10298
/bound_moiety="5-392-280.probe"
variation
10286
/note="5-392-280 : polymorphic base G or T"
primer_bind
10287..10305
/note="5-392-280.mis complement"
variation
10370
/note="5-392-364 : insertion of G"
primer_bind
10411..10430
/note="5-392.rp complement"
exon
26810..26897
/note="exon D"
exon
31357..31471
/note="exon E"
exon
34261..34404
/note="exon F"
exon
37377..37466
/note="exon S"
primer_bind
39556..39574
/note="4-59.rp"

exon
39704..40858
/note="exon T"
primer_bind
39877..39896
/note="4-58.rp"
primer_bind
39925..39943
/note="4-58-318.mis"
misc_binding
39932..39956
/bound_moiety="4-58-318.probe"
variation
39944
/note="4-58-318 : polymorphic base G or T"
primer_bind
39945..39963
/note="4-58-318.mis complement"
primer_bind
39953..39970
/note="4-59.pu complement"
primer_bind
39954..39972
/note="4-58-289.mis"
misc_binding
39961..39985
/bound_moiety="4-58-289.probe"
variation
39973
/note="4-58-289 : polymorphic base G or C"
primer_bind
39974..39992
/note="4-58-289.mis complement"
primer_bind
40242..40259
/note="4-58.pu complement"
primer_bind
41137..41154
/note="4-54.rp"
primer_bind
41366..41384
/note="4-54-199.mis"
misc_binding
41373..41397
/bound_moiety="4-54-199.probe"
primer_bind
41385..41403
/note="4-54-180.mis"
variation
41385
/note="4-54-199 : polymorphic base A or C"
primer_bind
41386..41404
/note="4-54-199.mis complement"
misc_binding
41392..41416
/bound_moiety="4-54-180.probe"
variation
41404
/note="4-54-180 : polymorphic base A or C"
primer_bind
41405..41423
/note="4-54-180.mis complement"
primer_bind
41564..41581
/note="4-54.pu complement"
primer_bind
42122..42141
/note="4-51.rp"
primer_bind
42213..42231
/note="4-51-312.mis"
misc_binding
42220..42244
/bound_moiety="4-51-312.probe"
variation
42232
/note="4-51-312 : polymorphic base G or C"
primer_bind
42233..42251
/note="4-51-312.mis complement"
primer_bind
42526..42543
/note="4-51.pu complement"
exon
50436..50545
/note="exon G"
primer_bind
67289..67309
/note="99-86.rp"
primer_bind
67456..67474
/note="99-86-266.mis"
misc_binding
67463..67487
/bound_moiety="99-86-266.probe"
variation
67475
/note="99-86-266 : polymorphic base A or C"
primer_bind
67476..67494
/note="99-86-266.mis complement"
primer_bind
67724..67741
/note="99-86.pu complement"
primer_bind
69182..69200
/note="4-88.rp"
primer_bind
69502..69520

```
misc_binding      /note="4-88-107.mis"
                  69509. .69533
bound_molecly="4-88-107.probe"
variation
                  /note="4-88-107 : polymorphic base A or G"
                  69521
primer_bind      /note="4-88-107 : polymorphic base A or G"
                  69522. .69540
primer_bind      /note="4-88-107.mis complement"
                  69609. .69626
primer_bind      /note="4-88.pu complement"
                  72698. .72715
primer_bind      /note="5-397.pu"
                  72819. .72837
misc_binding      /note="5-397-141.mis"
                  72826. .72850
variation         /bound_molecly="5-397-141.probe"
                  72838
primer_bind      /note="5-397-141 : polymorphic base G or T"
                  72839. .72857
exon             /note="5-397-141.mis complement"
                  72881. .72918
primer_bind      /note="exon H"
                  73099. .73117
primer_bind      /note="5-397.tp complement"
                  75858. .75877
exon             /note="5-398.pu"
                  75989. .76151
primer_bind      /note="exon I"
                  76041. .76059
misc_binding      /note="5-398-203.mis"
                  76048. .76072
variation         /bound_molecly="5-398-203.probe"
                  76060
primer_bind      /note="5-398-203 : polymorphic base A or C"
                  76061. .76079
primer_bind      /note="5-398-203.mis complement"
                  76289. .76306
primer_bind      /note="5-398.tp complement"
                  81006. .81025
primer_bind      /note="99-12738.pu"
                  81234. .81252
misc_binding      /note="99-12738-248.mis"
                  81241. .81265
variation         /bound_molecly="99-12738-248.probe"
```

Query	Match	Similarity	19.5%	Score 238	DB 6	Length 240825
Best Local	Similarity	64.3%	Pred. No. 2.9e-44			
Match	545	Conservative	0	Mismatches 265	Indels 38	Gaps 11
QY	221	CCCGAGAAACCCGTGATACAGTGAAGAAAGGATGTTGGTTCCTCGACATACCT	280			
Db	158572	CTCAAAAAAAAAAAAAAAAAAAAAAAAAAGTTGTGTTCTCTTTTACACGACACTT	158513			
QY	281	TCACTCAAATATCTATGTGTTACTTAA CAGCAA TTATCAGGCCAAACCTTAAGTTT	340			
Db	158512	AAATTCCTAATTTCTATTGTATGTATGATGAATCTATCAGTACAGTGATTTTGTAGAGT	158453			
QY	341	TATTTTATAGTATTTTGTGTTAAGCACAGACATGTGAGTCTGACAAAACGTGATTTG	400			
Db	158452	GAATATCCAGTGAGAGTTTAAAAAGA -ATAAGGTCAGAGCTCAG -AAATCGATGTTTT	158395			
QY	401	GTAACCTGATTTTATTA-----TATCAACTGGGTTTAAATTAATA-----AA	443			
Db	158394	CTAACGTGACTTAAATAACGATCTTCTCTATCCAACTGTTGGAAATGAAATCTTTGTGTA	158335			
QY	444	AATGTCATTACTTAAAAAAAAAACCAATACCAACAAGCTTTACTCCCTTGGAAA	503			
Db	158334	ACTTGCGTAGTGCTCAAGATCACTAAATCAAGAACTATTTTTCCCTCTGAAAAG	158275			
QY	504	GCACATTTTCAAGGG---CTGATCTTAGCCTTATATTTTTCATTAAGAAATTAACCA	560			
Db	158274	GCACACGTTTACAGAGCCAAACTTCTCTCTTTTATTTGTATTAACAAAATAAACC	158215			
QY	561	GGTCCCGATATAGCTGTAAATTTTATTCCTA-----AAAGAACAGAAATTTTCACGTATG	613			

DB 158214 AGTCCGACCTATTTGTAGTATTTTATTTCTTAAGGAAAAAACAAGAACTTTCATTGTGA 1581555

QY 614 CTTTAAATTAAGTATTACCTCAGATACCTCTGCAAGCTTGAACCAACGAGCAG-- 671

DB 158154 CTTCAACATTAAAGTTATTACCTCAGATATTTTCCAGCTTGACAGCGCAAAATACGTT 1580959

QY 672 ACAGACAACAGAGCCCCA-GCTACTCTTAGGAAA---TAAATTAGGTGTGCTCTGA 726

DB 158094 TCAGACAAAAGAGATCAACTGCTCTCTTAGGAAATATTAAATTGGGGGTGTGCTTAGGA 158035

QY 727 CATGCCAAGGGTCTTGTGCTGTGTCG-TGTTCCAGAAAGCTTTGCACTACACATC 785

DB 158034 AATGCCACAGGGTCTCTGTAAAGATCGGTTTTCCAGAGGGTTCTGACAGCATGGGTCC 1579757

QY 786 CTTTGGGCGAGTACACATATCTGTGATTTTCCGTGCTGCTGCTAGTGAAGCCCTAC 845

DB 157974 TGGTTGGAGAGCGACGATTTCTGCTGTGATTTTCCGTGCTGCTGATGACCCCTAC 1579151

QY 846 AGGAAGATAGTGGTGTAGCCAGG-GGGCGAGCGGCTGGCTGCACATGTTCTGGTCTCT 904

DB 157914 AGGAAGATTAACGGCTTAGCGCAGAGAGGGCGAGCAGCCCACTACACATGTTCTGGTCTCT 1578555

QY 905 TATCAACTTATATTAAGGGAAGGAAGTGAATTGTCGGTACTGACACCTGTAGACTC 964

DB 157854 TATCAACTTATATTAAGGGAAGGAAGTGAATTGTCGGTACTGACACCTGTAGAACTC 1577955

QY 965 AGGGAGAAACAAGAGTCCGTGCAGACCTCTGAGTGAAGCAGGGCTCTCTTCCTCTC 1024

DB 157794 TGGGAGAGAGGAACAAGAGACCGGTGAAGCTCTGTGAAGCTGACACAGCCCTCCC 1577355

QY 1025 AGGACAGC 1032

DB 157734 AAGTGAGC 157727

FEATURES	LOCATION/QUALIFIERS
LOCUS	AX523960
DEFINITION	Sequence 1 from Patent WO02066641.
ACCESSION	AX523960
VERSION	AX523960.1
KEYWORDS	GI:25168889
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 Barry, C. and Chumakov, I.
TITLE	Pg-3 and diallelic markers thereof
JOURNAL	Patent: WO 02066641-A 1 29-AUG-2002;
GENSET	(FR)
SOURCE	Location/Qualifiers
	1..240825
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
	1..2000
	/note="5' regulatory region"
	1823..1840
	/note="5'-390..pu"
	1980..1998
	/note="5'-390-177.mis"
	1987..2011
	/bound_moiety="5'-390-177.probe"
	1999
	/note="5'-390-177 : polymorphic base G or C"
	2000..2018
	/note="5'-390-177.mis complement"
	2001..2079
	/note="exon A"
	2108..2125

```
primer_bind /note="5-390.rp complement"
4559.4577
/note="5-391.pu"
primer_bind 4582.4600
/note="5-391-43.mis"
misc_binding 4589.4613
/bound_moiety="5-391-43.probe"
variation 4601
/note="5-391-43 : polymorphic base A or G"
primer_bind 4602.4620
/note="5-391-43.mis complement"
exon 4627.4718
/note="exon B"
primer_bind 4891.4908
/note="5-391.rp complement"
primer_bind 10007.10025
/note="5-392.pu"
exon 10115.10233
/note="exon C"
primer_bind 10209.10227
/note="5-392-222.mis"
misc_binding 10216.10240
/bound_moiety="5-392-222.probe"
variation 10228
/note="5-392-222 : polymorphic base G or T"
primer_bind 10229.10247
/note="5-392-222.mis complement"
primer_bind 10267.10285
/note="5-392-280.mis"
misc_binding 10274.10298
/bound_moiety="5-392-280.probe"
variation 10286
/note="5-392-280 : polymorphic base G or T"
primer_bind 10287.10305
/note="5-392-280.mis complement"
variation 10370
/note="5-392-364 : insertion of G"
primer_bind 10411.10430
/note="5-392.rp complement"
exon 26810.26897
/note="exon D"
exon 31357.31471
/note="exon E"
exon 34261.34404
/note="exon F"
exon 37377.37466
/note="exon S"
primer_bind 39556.39574
/note="4-59.rp"
exon 39704.40858
/note="exon T"
primer_bind 39877.39896
/note="4-58.rp"
primer_bind 39925.39943
/note="4-58-318.mis"
misc_binding 39932.39956
/bound_moiety="4-58-318.probe"
variation 39944
/note="4-58-318 : polymorphic base G or T"
primer_bind 39945.39963
/note="4-58-318.mis complement"
primer_bind 39953.39970
/note="4-59.pu complement"
primer_bind 39954.39972
/note="4-58-289.mis"
misc_binding 39961.39985
/bound_moiety="4-58-289.probe"
variation 39973
/note="4-58-289 : polymorphic base G or C"
primer_bind 39974.39992
/note="4-58-289.mis complement"
40242.40259
/note="4-58.pu complement"

primer_bind 41137.41154
/note="4-54.rp"
primer_bind 41366.41384
/note="4-54-199.mis"
misc_binding 41373.41397
/bound_moiety="4-54-199.probe"
primer_bind 41385.41403
/note="4-54-180.mis"
variation 41385
/note="4-54-199 : polymorphic base A or C"
primer_bind 41386.41404
/note="4-54-199.mis complement"
misc_binding 41392.41416
/bound_moiety="4-54-180.probe"
variation 41404
/note="4-54-180 : polymorphic base A or C"
primer_bind 41405.41423
/note="4-54-180.mis complement"
primer_bind 41564.41581
/note="4-54.pu complement"
primer_bind 42122.42141
/note="4-51.rp"
primer_bind 42213.42231
/note="4-51-312.mis"
misc_binding 42220.42244
/bound_moiety="4-51-312.probe"
variation 42232
/note="4-51-312 : polymorphic base G or C"
42233.42251
/note="4-51-312.mis complement"
42526.42543
/note="4-51.pu complement"
exon 50436.50545
/note="exon G"
67289.67309
/note="99-86.rp"
67456.67474
/note="99-86-266.mis"
67463.67487
/bound_moiety="99-86-266.probe"
67475
/note="99-86-266 : polymorphic base A or G"
67476.67494
/note="99-86-266.mis complement"
67724.67741
/note="99-86.pu complement"
69182.69200
/note="4-88.rp"
69502.69520
/note="4-88-107.mis"
69509.69533
/bound_moiety="4-88-107.probe"
69521
/note="4-88-107 : polymorphic base A or G"
69522.69540
/note="4-88-107.mis complement"
69609.69626
/note="4-88.pu complement"
72698.72715
/note="5-397.pu"
72819.72837
/note="5-397-141.mis"
72826.72850
/bound_moiety="5-397-141.probe"
72838
/note="5-397-141 : polymorphic base G or T"
72839.72857
/note="5-397-141.mis complement"
72861.72918
/note="exon H"
73099.73117
/note="5-397.rp complement"
75858.75877
```



```

QY      304 TACAGACAATTATAGGCGCAACACTTAAAGTTTATTTGTAAGTATTTTGCTTA 363
      3916 TGTAAAGAACTTAATAGTACAGCTGATTTTATAGCTGATATCAAGTGGAGCTTTT 3975
QY      364 AGGACAGACATGTGAGTCTGAGAAAAGTATGTTGGTAACTGATTAATA----- 417
      3976 AAGA-ATAAGTTCAGAGCTGAG-ATACTGATGTTTCTTACGTGACTTATATACGATC 4033
QY      418 ---TATCAACTGCGGTTAAATATAA-----AAATGTGCATTACTTAAAAA 466
      4034 TTTCTATCCAACTGTCGGAATGAAATCTTTTGTAACTGTGCTAGTGGTCAAGATC 4093
QY      467 ACCAATACCAACAAGACTTACTTCCCTTGAAAAAGCACATTACAGAG--CTGAT 523
      4094 AACTAAATCAAGAACTATATTTTCCCTGTGAAGGACAGGTTTACAGAGGCCAAGC 4153
QY      524 CTAGCCTTATATTTATCAATAGAAAATTAACCAAGGCTCCGATATAGCTTAATTT 583
      4154 TTTCTTCTCTTTATTTTGTATTAACAAAATTAACCAAGTCTGACCTATTTGTATTT 4213
QY      584 ATTCTA-----AAGAAGACAGAACTTCACTATGCTTTAAATTAAGTATTA 636
      4214 TTTATCTCTAAGGAAAAAAGAGAACTTTCATTTGTAATTCATTAAGTATTA 4273
QY      637 CAGATCTCTGCAAGCTTACCTTACAAAGAGAG--ACAGACACAGAGCCCCA-GCTA 693
      4274 CAGATATTTTGGCCAGCTTACGACAGGCAAAATAGTTTACAGAAAAGATCACTGCT 4333
QY      694 CTCTCTAGAGAA-----TAAATAGAGTGGTGGCTCTGACATGCCAGAGGGCTTGGCTG 749
      4334 CTCTCTAGAAATACCTTAATTTGGAGTGGTGGCTTGAAGAAATGCCAGAGGGCTGTAACG 4393
QY      750 GTCTGTG-TTCCAGAGAGGCTTCTGACATACAGATCTTTGGGAGCAGTAAGCACTATGC 808
      4394 ATCGGTTTTTCCAGAGAGGTTTCTGACAGATGGGTCCTGGTTGAGAGGAGGATTTCTGC 4453
QY      809 TCTGATTTTTCCTGTTGCTGCTGCTAGTACCCCTTACAGAGATATGTGGGTGAGCCAG 868
      4454 TCTGATTTTTCCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4513
QY      869 -GGGCGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 927
      4514 AGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4573
QY      928 GGAAGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 987
      4574 GGAAGATGATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4633
QY      988 CAGACTCTGAGAGTGAAGAGGCTGCTCTTCTCTCAGAGACGC 1032
      4634 GTGAAGCTGCTCTGTMAAAGCTGACACAGCCCTCCCAAGTGC 4678

```

```

RESULT 9
LOCUS   BC027216          2471 bp      mRNA      linear      ROD 04-OCT-2003
DEFINITION Mus musculus angiotensin 2, mRNA (cDNA clone MGC:25292
IMAGE:3494566), complete cds.
ACCESSION BC027216
VERSION   BC027216.1 GI:20070843
KEYWORDS MGC.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 2471)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Martusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Cavaletto,T.L.,

```

```

REMARK
COMMENT
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
2 (bases 1 to 2471)
Strausberg,R.
Submitted (04-Apr-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgdbcm.tmc.edu
Gundaracine, P.H., Garcia, A.M., Lu, X., Huliy, S.W., Louised, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINT at: http://image.llnl.gov
Series: IRAX Plate: 30 Row: b Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 667156.
Location/Qualifiers
1..2471
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Czech 11"
/db_xref="taxon:10090"
/clone="MGC:25292 IMAGE:3494566"
/tissue_type="Mammary tumor metastasized to lung. Tumor
arose spontaneously from a senescent normal mammary
(clonal) outgrowth infected with the virus MMTV."
/clone_1fb="NCI CGAP_Lu29"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..2471
/gene="Agrp2"
/note="synonyms: Ang2, Ang-2"
/db_xref="GeneID:11601"
/db_xref="MGI:1202890"
238..1728
/gene="Agrp2"
/codon_start=1
/product="Agrp2 protein"
/protein_id="AAH27216.1"
/db_xref="GI:20070844"
/db_xref="GeneID:11601"
/db_xref="MGI:1202890"
/translatability="WMOIIFLFGMDLVLAASVSNFRKSVSTGRQYOVQVQVPCSYT
FLPFTDSSSSSPYMSNAVORDADVDSDVQVLENTLENTLENTLENTLENTLENTLENT
DNMKEMVEIQONVQNTAVAVIEIGTSLNQTAKRTLDVAVQVAVNQTTRLEIOL
LOHSTINKLEKQIIDQTSINKLONKNSFLEQKVLDMEGKHSBOLQSNKEXDELOV

```

CDS

gene

FEATURES

source

Kowis, C., Kraft, C.L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lohenschuh, L., Louie, H., Lozano, R., Lu, X., Ma, J., Maheshwari, M., Mahindratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenan, E., Mitoavajevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okunribido, G., Olarnpungsoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L., Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soes, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabors, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, I., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.

Unpublished
2 (bases 1 to 221218)
Direct Subassembly
Submitted (16-NOV-2004) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 221218)
Cow Genome Sequencing Consortium.
Direct Subassembly
Submitted (01-JUL-2005) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Jun 26, 2005 this sequence version replaced gi:58037895. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: FANL
Center clone name: CH240-8821
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 213012 bases at least Q40
Consensus quality: 215322 bases at least Q30
Consensus quality: 217444 bases at least Q20
Estimated insert size: 214243; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently

consists of 27 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
2088: contig of 2088 bp in length
2089
2138: gap of 50 bp
2139
11896: contig of 9758 bp in length
11897
11946: gap of 50 bp
11947
13770: contig of 1824 bp in length
13771
13820: gap of 50 bp
13821
17181: contig of 3361 bp in length
17231
17231: gap of 50 bp
17232
18621: contig of 1390 bp in length
18671
18671: gap of 50 bp
18672
18672: contig of 43144 bp in length
61816
61816: gap of 50 bp
61865: gap of 50 bp
65984: contig of 4119 bp in length
65985
66034: gap of 50 bp
66035
87544: contig of 21510 bp in length
87545
87545: gap of 50 bp
87593
95737: contig of 8143 bp in length
95738
95738: gap of 50 bp
95788
100716: contig of 4929 bp in length
100717
100766: gap of 50 bp
100767
108628: contig of 7862 bp in length
108629
108678: gap of 50 bp
108679
109886: contig of 1158 bp in length
109887
119073: contig of 9187 bp in length
119074
119123: gap of 50 bp
119124
127103: contig of 7980 bp in length
127104
127153: gap of 50 bp
127154
134419: contig of 7266 bp in length
134420
134469: gap of 50 bp
155171
155171: contig of 20702 bp in length
155221
155221: gap of 50 bp
190589
190589: contig of 35368 bp in length
190590
190639: gap of 50 bp
201726
201726: contig of 11087 bp in length
201826
201826: gap of unknown length
203687
203687: contig of 1861 bp in length
203737
203737: gap of 50 bp
211192
211192: contig of 7455 bp in length
211193
212346: gap of unknown length
212347
212346: contig of 1054 bp in length
212347
212446: gap of unknown length
212447
213466: contig of 1020 bp in length
213467
213566: gap of unknown length
213567
214871: contig of 1305 bp in length
214872
214971: gap of unknown length
214972
214972: contig of 1227 bp in length
216199
216199: gap of unknown length
216298
216298: gap of unknown length
218080
218080: contig of 1782 bp in length
218081
218180: gap of unknown length
220028
220028: contig of 1848 bp in length
220029
220128: gap of unknown length
220129
221218: contig of 1090 bp in length.
220129
Location/Qualifiers
1..221218
/organism="bos taurus"
/mol_type="genomic DNA"
/db_xref="taxon:9913"
/clone="CH240-8821"
2089..2138
/estimated_length=50
11897..11946
/estimated_length=50
13771..13820
/estimated_length=50
17182..17231

FEATURES
source
gap
gap
gap
gap

gap /estimated length=50
 18622. .18671 /estimated length=50
 gap 61816. .61865 /estimated length=50
 gap 65985. .66034 /estimated length=50
 gap 87545. .87594 /estimated length=50
 gap 95738. .95787 /estimated length=50
 gap 100717. .100766 /estimated length=50
 gap 108629. .108678 /estimated length=50
 gap 109837. .109886 /estimated length=50
 gap 119074. .119123 /estimated length=50
 gap 127104. .127153 /estimated length=50

Query Match 15.6%; Score 190.6; DB 14; Length 221218;
 Best Local Similarity 67.6%; Pred. No. 2.4e-33;
 Matches 390; Conservative 0; Mismatches 169; Indels 18; Gaps 8;

QY 432 AAAAAAATGCTAATAAAAAAACAATACCAACAGACTTACTT 491
 DB 141965 AGTTTGAAGTCTGCTAAAGTTCAAAAATCAAGTAATCCAAAGAACTATTTC 141906
 QY 492 CCCCTTGAA--AAGCATTACAAAGGCTGATCTTAACTTATTTTCAATAAAG 548
 DB 141905 CCAATGAAAGAGCAGAGTTTCAAGAGGAGACTTCTTCTTTGTTTGAATAAAC 141846
 QY 549 AAAAAATCAAGGCTCCGATATA--GCTTAATTTTATTCCTAAAGAGAACTT 605
 DB 141845 AAAATTAACCAAGCTCTGACATATTTGACGATTTTATTCCTAAAGAAA--AAACAG 141788
 QY 606 TCATGCTTAAATTAAGGATTAAGTACTGATGATCTGCAAGCTTACCTTAAAC 665
 DB 141787 GAATTTTCAATTTAAAGCTATTGCCCCAGATTAATACAGCTTGGCCTGTGAAA 141728
 QY 666 GAGCAGACAGACAAAGAGCCCAAGCTACTTCT--AGGAATTAATTAAGGCTG 719
 DB 141727 TCAGTTTCAAGACAGAGAGACCTTACAGGCTCTCTAGGAAACAGTCAATTAAGGCTG 141668
 QY 720 CCTTGAACATGCCAGGCTTTGCTGCTGTG--TTCCAGAAAGCTTCTGCACTA 778
 DB 141667 CCTGAAATACCCAGGCTCTGTAACTGATCGGTTTTCCAGAGGCTTCTGACAGA 141608
 QY 779 CACAGT-CCTTTGGGAGTAAAGCACTATGCTGATTTTCTGTTGCTGGCTAGTGA 837
 DB 141607 TTAGTCCCGGTGAGCACTACATCAAGCCCGATTTTCTGCTGGCTGGTGA 141548
 QY 838 -CCCCCTTACAGAAAGTATGAGTGAAGCCAGG-GGGCGAGAGCGCTGCTGACATGCT 895
 DB 141547 CCCCCCTTACAGAAAGTAAAGGCAAGCCGAGAGGCGGAGCGGCTGACATGCT 141488
 QY 896 GGCTGCTTATCACTTATCATTAAGGAAAGAAAGTATTTGAGTACTGACAC 955
 DB 141487 GCTGCTCTTATCACTTATCATTAAGGAAAGAAAGTATTTGAGTACTGACAC 141428
 QY 956 TGTAGACTGAGGAGAAACAAGAGTCCGTGACAGC 992
 DB 141427 TGCAGATCTGGGAGAGAGCAACAAGCAGCGCTAAC 141391

RESULT 12
 HSI64G8F
 LOCUS HSI64G8F 290 bp DNA linear PRI 18-OCT-1995
 DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 164g8,
 forward read cpg164g8.ft1a.
 ACCESSION 257119
 VERSION 257119.1 GI:1028350

KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Cpg island; genomic MseI fragment.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.

REFERENCE
 AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
 TITLE Purification of Cpg islands using a methylated DNA binding column
 JOURNAL Nat. Genet. 6 (3), 236-244 (1994)
 PUBMED 801384

REFERENCE
 AUTHORS Dodsworth,S.J., Huckle,E., Wilkinson,P. and Micklem,G.
 TITLE Direct Submision
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, England. E-mail contact: humquery@sanger.ac.uk

COMMENT
 Vector: pGEM-5Zf(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
 http://www.hgmp.mrc.ac.uk/ for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source location/Qualifiers

1..290
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="164g8"
 /sex="male"
 /library_type="blood"
 /clone_lib="CGI-1"
 /dev_stage="adult"

ORIGIN

Query Match 14.4%; Score 176.2; DB 8; Length 290;
 Best Local Similarity 81.9%; Pred. No. 7.5e-30;
 Matches 227; Conservative 0; Mismatches 48; Indels 2; Gaps 2;

QY 706 TAAATGAGTGTGCTCTGACATGCCAGGAGCTTGTGCTGCTGTG--TTCCAGA 764
 DB 1 TAAATGAGTGTGCTCTGACATGCCAGGAGCTTGTGCTGCTGTG--TTCCAGA 60
 QY 765 AGGCTTGCAGTACACAGTCTTTGGGGAGTAAAGCACTATGCTGATTTTCCGT 824
 DB 61 GGGTTTTCGACATGGGCTCTGTTGAGGAGGAGGATTCGTCTGATTTTTCCTG 120
 QY 825 GCTGCTAGTGAACCCCTTACAGAAAGATAGTGGTGAAGCCAGG-GGGCGAGCGGCTG 883
 DB 121 GCTGGGTAGTGAACCCCTTACAGAAAGATAGTGGTGAAGCCAGGCGGAGCAGCCA 180
 QY 884 CTGACATGTTGCTGCTCTTATCACTTATCATTAAGGAAAGAAAGTATGATTC 943
 DB 181 CTACACATGTTGCTGCTCTTATCACTTATCATTAAGGAAAGAAAGTATGATTC 240
 QY 944 GGATACAGACCTGATGCTGAGGAGGAGAAACAAGA 980
 DB 241 GGATACAGACCTGATGCTGAGGAGGAGAGAAACA 277

RESULT 13
 AC141003
 LOCUS AC141003 132392 bp DNA linear HTG 27-MAR-2003
 DEFINITION Rattus norvegicus clone CH230-526115, *** SEQUENCING IN PROGRESS
 AC141003
 ACCESSION AC141003.1 GI:28875862
 VERSION HTG; HTGS PHASE1.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Muridae; Murinae; Rattus.
 1 (bases 1 to 132392)
 Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,

Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benhmed, F., Bialdo, K., Blair, J., Blankenburg, K., Blych, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Ledow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensbawa, L., Loulesged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, B., Mawhney, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokeleneh, O., Okunolu, G., Olarunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Yoon, V., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

TITLE
Direct Submission

Unpublished
2 (bases 1 to 132392)

REFERENCE
AUTHORS

Worley, K. C.

TITLE
Direct Submission

Submitted (07-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 132392)

REFERENCE
AUTHORS
TITLE
JOURNAL

Worley, K. C.

Submitted (27-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GKXS

Center clone name: CH230-526115

----- Summary Statistics

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 98530 bases at least Q40
Consensus quality: 105678 bases at least Q30
Consensus quality: 110765 bases at least Q20
Estimated insert size: 86014; sum-of-contigs estimation
Quality coverage: 1x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Gendank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 64 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1018:	contig of 1018 bp in length
1019	1118:	gap of unknown length
1119	2164:	contig of 1046 bp in length
2165	2264:	gap of unknown length
2265	3340:	contig of 1076 bp in length
3341	3440:	gap of unknown length
3441	4445:	contig of 1005 bp in length
4446	4545:	gap of unknown length
4546	5558:	contig of 1013 bp in length
5559	5658:	gap of unknown length
5659	6671:	contig of 1013 bp in length
6672	6771:	gap of unknown length
6772	7849:	contig of 1078 bp in length
7850	7949:	gap of unknown length
7950	9069:	contig of 1120 bp in length
9070	9169:	gap of unknown length
9170	10301:	contig of 1132 bp in length
10302	10401:	gap of unknown length
10402	11937:	contig of 1536 bp in length
11938	12037:	gap of unknown length
12038	13406:	contig of 1369 bp in length
13407	13506:	gap of unknown length
13507	14990:	contig of 1484 bp in length
14991	15090:	gap of unknown length
15091	16107:	contig of 1017 bp in length
16108	16207:	gap of unknown length
16208	17285:	contig of 1078 bp in length
17286	17385:	gap of unknown length
17386	18963:	contig of 1578 bp in length
18964	19063:	gap of unknown length
19064	20794:	contig of 1731 bp in length
20795	20894:	gap of unknown length
20895	22416:	contig of 1522 bp in length
22417	22516:	gap of unknown length
22517	23639:	contig of 1123 bp in length
23640	23739:	gap of unknown length
23740	24842:	contig of 1103 bp in length
24843	24942:	gap of unknown length
24943	26028:	contig of 1066 bp in length
26029	26128:	gap of unknown length
26129	27718:	contig of 1590 bp in length
27719	27818:	gap of unknown length
27819	29221:	contig of 1403 bp in length
29222	29321:	gap of unknown length
29322	30448:	contig of 1127 bp in length
30449	30548:	gap of unknown length
30549	31894:	contig of 1346 bp in length
31895	31994:	gap of unknown length
31995	33378:	contig of 1384 bp in length
33379	33478:	gap of unknown length
33479	34875:	contig of 1357 bp in length
34876	34975:	gap of unknown length
34976	36002:	contig of 1027 bp in length
36003	36102:	gap of unknown length
36103	37499:	contig of 1397 bp in length
37500	37599:	gap of unknown length
37600	39671:	contig of 2072 bp in length

CC The present sequence is that of the cis-acting regulatory region of the
 CC murine Ang-2 gene located 5' to the start of transcription. This cis-
 CC acting DNA regulates preferential expression in endothelial muscle cells
 CC of a polypeptide-encoding DNA to which it is operably linked. The
 CC invention provides compositions that include a polynucleotide encoding a
 CC reporter gene (e.g. a light-generating moiety), a polynucleotide encoding a
 CC selectable marker (e.g. an antibiotic) and optionally a regulatory
 CC element (e.g. the Ang-2 cis-acting regulatory sequence, or a portion of
 CC it that retains the ability to effect transcription of operably linked
 CC sequences in endothelial cells). The composition is used for imaging
 CC cells. The cells may be healthy cells or tumor cells, in which case
 CC tumour growth, engraftment and/or metastasis can be monitored by
 CC observing light emitted from the light-generating gene product. Also
 CC provided are vectors comprising the polynucleotides, cells (e.g. primary
 CC cells, cells from established cell lines, or tumor cells) comprising the
 CC vector or polynucleotides, and transgenic animals and methods for their
 CC production. The transgenic animals are used in methods for the
 CC identification of a compound capable of modulating angiogenesis or which
 CC has tumour inhibitor activity, or which is a modulator of a signal
 CC transduction pathway.

XX Sequence 1220 BP; 370 A; 254 C; 295 G; 301 T; 0 U; 0 Other;

Query Match 100.0%; Score 1220; DB 12; Length 1220;
 Best Local Similarity 100.0%; Pred. No. 1e-312;

Matches 1220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCTTTCCAGTACCGATCTCTGACAGATTAACTTCTAGTACATGAAAGGGGTGTGATCTC 60
 Db 1 GAGCTTTCCAGTACCGATCTCTGACAGATTAACTTCTAGTACATGAAAGGGGTGTGATCTC 60
 QY 61 TGGACGACGAGACCAAGAGCTGGAAGTGTGTTAGAGTCAAGTGAAGCCCGACGTTT 120
 Db 61 TGGACGACGAGACCAAGAGCTGGAAGTGTGTTAGAGTCAAGTGAAGCCCGACGTTT 120
 QY 121 ATGGCCAGGGGCTTTGAACTTAATTAAGGGGAAAGTATTGCTGAGCCCACTGAC 180
 Db 121 ATGGCCAGGGGCTTTGAACTTAATTAAGGGGAAAGTATTGCTGAGCCCACTGAC 180
 QY 181 TGGGACTAATTTAATGAGGAATGCGACAGAGTGTGAGCCCGAGAAACCTGTATCA 240
 Db 181 TGGGACTAATTTAATGAGGAATGCGACAGAGTGTGAGCCCGAGAAACCTGTATCA 240
 QY 241 GTGAGGAAAGGTGTATGTTTCTCTGACATACCTCACTCAATATCTATTGTT 300
 Db 241 GTGAGGAAAGGTGTATGTTTCTCTGACATACCTCACTCAATATCTATTGTT 300
 QY 301 ACTTAACAGACAATTAATGAGCCAAACCACTTAAGTTTATTTGATATGATTTTGTG 360
 Db 301 ACTTAACAGACAATTAATGAGCCAAACCACTTAAGTTTATTTGATATGATTTTGTG 360
 QY 361 TTAAGGCACAGACATGTGAGTCTGAGAAACTGATGTGTAACCTGATTAATATAT 420
 Db 361 TTAAGGCACAGACATGTGAGTCTGAGAAACTGATGTGTAACCTGATTAATATAT 420
 QY 421 CAAACCTGGTTAAATTAATTAATTAATGTCATTAATTAATTAATTAATTAATTA 480
 Db 421 CAAACCTGGTTAAATTAATTAATTAATGTCATTAATTAATTAATTAATTAATTA 480
 QY 481 AGACTTTAATCTCCCTTGGAAAGACATTTACAGAGGCTGATCTTATCTTTATTTA 540
 Db 481 AGACTTTAATCTCCCTTGGAAAGACATTTACAGAGGCTGATCTTATCTTTATTTA 540
 QY 541 CAATTAAGAAATAAACAAGGTCCGATATAGCTGTATTTATCTTAAGAAGACAGA 600
 Db 541 CAATTAAGAAATAAACAAGGTCCGATATAGCTGTATTTATCTTAAGAAGACAGA 600
 QY 601 AACTTCACTATGCTTTAAATTAAGTATTAATCTCAATCTCTGCAAGCTTAGCCTA 660
 Db 601 AACTTCACTATGCTTTAAATTAAGTATTAATCTCAATCTCTGCAAGCTTAGCCTA 660
 QY 661 CAAACGACGACAGACAAAGAGCCGAGCTACTCTAGAAATTAATTAGGGTGTGTC 720
 Db 661 CAAACGACGACAGACAAAGAGCCGAGCTACTCTAGAAATTAATTAGGGTGTGTC 720

Db 661 CAAACGACGACAGACAAAGAGCCCGAGCTACTCTAGAAATTAATTAGGGTGTGTC 720
 QY 721 CTCTGACATGCCCGAGGGGCTTTGTGGCTGTGTGTGTTCAGAAAGGCTTGTGACATCA 780
 Db 721 CTCTGACATGCCCGAGGGGCTTTGTGGCTGTGTGTGTTCAGAAAGGCTTGTGACATCA 780
 QY 781 CAGTCTTTGGGGCAGTAAAGCACTATGCTGTGATTTTCTGTGCTGGCTGACATGACCC 840
 Db 781 CAGTCTTTGGGGCAGTAAAGCACTATGCTGTGATTTTCTGTGCTGGCTGACATGACCC 840
 QY 841 CCTACAGGAATAGTGGGTGAGCCAGGGGCGAGCGGCTGCTGCAATGTCTGGCTG 900
 Db 841 CCTACAGGAATAGTGGGTGAGCCAGGGGCGAGCGGCTGCTGCAATGTCTGGCTG 900
 QY 901 CTCTTATCACTTATCATATTAAGGAAGAAAGTATGATTTCCGATCTGACACTGTAG 960
 Db 901 CTCTTATCACTTATCATATTAAGGAAGAAAGTATGATTTCCGATCTGACACTGTAG 960
 QY 961 ACTCAGGGGAGAAACAAAGAGTCCGTGAGACCTCTGAGTGAAGCAGGGCTGCTCTTCC 1020
 Db 961 ACTCAGGGGAGAAACAAAGAGTCCGTGAGACCTCTGAGTGAAGCAGGGCTGCTCTTCC 1020
 QY 1021 TCTCAGACAGCTCCGATGTGCGGGGAGAAAGAAAGAGACAGACCTGGGAAA 1080
 Db 1021 TCTCAGACAGCTCCGATGTGCGGGGAGAAAGAAAGAGACAGACCTGGGAAA 1080
 QY 1081 GAGCCTGTGCGGGACGAGAAAGGCTCTCATGATGACTTATTCACACGGCACAGCCCT 1140
 Db 1081 GAGCCTGTGCGGGACGAGAAAGGCTCTCATGATGACTTATTCACACGGCACAGCCCT 1140
 QY 1141 GTGCTTTAGACAGACGCTGAGAGCTCAGAGCGAAAGTTTGCTGAATCTCACAGTTTAAC 1200
 Db 1141 GTGCTTTAGACAGACGCTGAGAGCTCAGAGCGAAAGTTTGCTGAATCTCACAGTTTAAC 1200
 QY 1201 CCAAAAAGAGAGAGAGATG 1220
 Db 1201 CCAAAAAGAGAGAGATG 1220

RESULT 2
 ID AD261847
 ID AD261847 standard; DNA, 2475 BP.
 XX AD261847;
 XX AC
 XX 30-JUN-2005 (first entry)
 DT
 DE Murine Asp22 gene, SEQ ID 1571.
 XX
 KW Drug screening; Peroxisome Proliferator-Activated Receptor gamma;
 KW PPAR gamma; gene; ds.
 XX
 OS Mus musculus.
 XX
 FN US2005084872-A1.
 PD 21-Apr-2005.
 XX
 XX 23-JAN-2004; 2004US-00764420.
 PF 24-JAN-2003; 2003US-0442797P.
 PR 30-MAY-2003; 2003US-0474413P.
 XX
 PA (LUMP/) LUM P Y.
 PA (TANV/) TAN Y.
 PA (DAIH/) DAI H.
 PA (MUIS/) MUISE E S.
 PA (BERG/) BERGER J P.
 PA (THOM/) THOMPSON J R.
 XX
 PI Lum PY, Tan Y, Dai H, Muise ES, Berger JP, Thompson JR;
 XX WPI, 2005-313963/32.
 DR


```
Db 158032 AATGCCAGGGGCTCTGTAAAGATCGGTTTTTCCAGAGGGTTCTGCAGCATGGGATCC 157973
Qy 786 CTTGGGGGAGTAAGCACTATGCTGTATTTTCCGTGTGCTAGAGACCCCTAC 845
Db 157972 TGGTTGAGGGGAGGCACTTCGCTGTATTTTCTGTGCTAGAGACCCCTAC 157913
Qy 846 AGGAAGATAGTGGGTGACCCAGG-GGGCGAGCGGCTGGCTGCACATGTCTGCTGCTCT 904
Db 157912 AGGAAGATAGCGGCTAAGCCAGAGGGCGAGCACCACTACATGTCTGCTGCTCT 157853
Qy 905 TATCAACTTATCATATATAAGGAAGAAAGTATGATTCGATCTGACACTGTAGACTC 964
Db 157852 TATCAACTTATCATATATAAGGAAGAAAGTATGATTCGATCTGACACTGTAGACTC 157793
Qy 965 AGGGAGAAACAAAGAGTCCGTCAGACCTCTGAGTAGACAGGCTGCTCTTCCCTTC 1024
Db 157792 TGGGGAGAGAGAACAAAGACCGTGAAGCTGCTGTAAAGCTGACACAGCCCTCCC 157733
Qy 1025 AGGACAGC 1032
Db 157732 AAGTGAGC 157725
```

```
RESULT 4
AAF24497/c
ID AAF24497 standard; cDNA; 240825 BP.
XX
XX AAF24497;
XX
XX 23-MAY-2001 (first entry)
XX
XX Human PG-3 gene.
XX
XX Human; PG-3; cancer; BRCA1; chromosome 8p23; ds.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FT misc_feature 1..2000
XX FT /*tag= a
XX FT /*note= "5' regulatory region"
XX FT primer_bind 1823..1840
XX FT /*tag= b
XX FT primer_bind 1980..1998
XX FT /*tag= c
XX FT misc_binding 1987..2011
XX FT /*tag= d
XX FT /*note= "binds probe"
XX FT allele replace(1999,C)
XX FT primer_bind /*tag= e
XX FT CDS complement(2000..2018)
XX FT /*tag= f
XX FT 2001..238825
XX FT /*tag= g
XX FT /*product= "PG-3"
XX FT /*note= "this sequence contains introns"
XX FT exon 2001..2079
XX FT /*tag= h
XX FT primer_bind /*label= A
XX FT 2108..2125
XX FT /*tag= i
XX FT primer_bind 4559..4577
XX FT /*tag= j
XX FT primer_bind 4582..4600
XX FT /*tag= k
XX FT misc_binding 4589..4613
XX FT /*tag= l
XX FT /*note= "binds probe"
XX FT allele replace(4601,G)
XX FT primer_bind /*tag= m
XX FT primer_bind complement(4602..4620)
XX FT /*tag= n
```

```
FT exon 4627..4718
FT /*tag= o
FT /*label= B
FT primer_bind 4891..4908
FT /*tag= p
FT primer_bind 10007..10025
FT /*tag= q
FT exon 10115..10233
FT /*tag= r
FT primer_bind 10209..10227
FT /*tag= s
FT misc_binding 10216..10240
FT /*tag= t
FT /*note= "binds probe"
FT allele replace(10228,T)
FT /*tag= u
FT primer_bind complement(10229..10247)
FT /*tag= v
FT primer_bind 10267..10285
FT /*tag= w
FT misc_binding 10274..10298
FT /*tag= x
FT /*note= "binds probe"
FT allele replace(10286,T)
FT primer_bind /*tag= y
FT complement(10287..10305)
FT /*tag= z
FT allele replace(10370,n)
FT /*tag= aa
FT primer_bind 10411..10430
FT /*tag= ab
FT exon 26810..26897
FT /*tag= ac
FT /*label= D
FT exon 31357..31471
FT /*tag= ad
FT /*label= E
FT exon 34261..34404
FT /*tag= ae
FT /*label= F
FT exon 37377..37466
FT /*tag= af
FT primer_bind /*label= S
FT 39556..39574
FT /*tag= ah
FT exon 39704..40858
FT /*tag= ag
FT /*label= T
FT primer_bind 39877..39896
FT /*tag= ai
FT primer_bind 39925..39943
FT /*tag= aj
FT misc_binding 39932..39956
FT /*tag= ak
FT /*note= "binds probe"
FT allele replace(39944,T)
FT /*tag= al
FT primer_bind complement(39945..39963)
FT /*tag= am
FT primer_bind 39953..39970
FT /*tag= an
FT primer_bind 39954..39972
FT /*tag= ao
FT misc_binding 39961..39985
FT /*tag= ap
FT /*note= "binds probe"
FT allele replace(39973,C)
FT primer_bind /*tag= aq
FT primer_bind complement(39974..39992)
FT /*tag= ar
FT primer_bind 40242..40259
FT /*tag= as
```

```
FT primer_bind 41137..41154
FT /*tag= ac
FT primer_bind 41366..41384
FT /*tag= au
FT misc_binding 41373..71397
FT /*tag= av
FT primer_bind 41385..41403
FT /*tag= ax
FT allele replace(41385,C)
FT primer_bind 41392..41416
FT /*tag= ay
FT misc_binding 41392..41416
FT /*tag= az
FT allele replace(41404,C)
FT primer_bind 41405..41423
FT /*tag= bd
FT primer_bind 41564..41581
FT /*tag= bc
FT primer_bind 42122..42141
FT /*tag= bd
FT primer_bind 42213..42231
FT /*tag= be
FT misc_binding 42220..42244
FT /*tag= bf
FT allele replace(42232,C)
FT primer_bind 42233..42251
FT /*tag= bh
FT primer_bind 42526..42543
FT /*tag= bi
FT exon 50436..50545
FT /*tag= bj
FT primer_bind 67289..67309
FT /*tag= bk
FT primer_bind 67456..67474
FT /*tag= bl
FT misc_binding 67463..67487
FT /*tag= bm
FT allele replace(67475,G)
FT primer_bind 67476..67494
FT /*tag= bn
FT primer_bind 67724..67741
FT /*tag= bp
FT primer_bind 69182..69200
FT /*tag= bq
FT primer_bind 69502..69520
FT /*tag= br
FT misc_binding 69509..69533
FT /*tag= bs
FT allele replace(69521,G)
FT primer_bind 69522..69540
FT /*tag= bt
FT primer_bind 69609..69626
FT /*tag= bu
FT primer_bind 72698..72715
FT /*tag= bv
FT primer_bind 72819..72837
FT /*tag= bw
FT misc_binding 72826..72850
FT /*tag= bx
FT allele replace(72838,T)
FT /*tag= by
```

```
FT primer_bind complement(72839..72857)
FT /*tag= ca
FT exon 72881..72918
FT /*tag= cb
FT label= H
FT primer_bind 73099..73117
FT /*tag= cc
```

Query Match 19.5%; Score 238; DB 4; Length 240825;
Best Local Similarity 64.3%; Pred. No. 2.1e-51;
Matches 545; Conservative 0; Mismatches 265; Indels 38; Gaps 11;

```
QY 221 CCCGAGAAACCTGATACAGTGAAGAAAGTGATGTTGTTCTCTCGACATCT 280
DB 158572 CTCAAAAAAAAAAAAAAAAAAAAAAAAAGTTGTATTCCTTCTACACACACTT 158513
QY 281 TCACCTCAATATCTATGTTACTTACAGCAATTAATAGGCCAACCACTTAAGTTT 340
DB 158512 AAATTTCTAATTTCTCTATGTAATGTAATGAATTAATCAAGTGAATTTTGAAGT 158453
QY 341 TATTTGTATAGTATTTTGTGTTAAGCACAGACATGTGAGTGTGAGAAATGATGTG 400
DB 158452 GAATATCCAAAGTGAAGTATTTTAAAGA-ATAAGGTCAAGACTAG-AACTGATGTTT 158395
QY 401 GTTACTTGATTTAATA-----TATCAACTGGGTTAAATAAAA-----AA 443
DB 158394 CTAACTGACTTAATTAACAGATCTTCTATCCAACTTCGGAAATGAATCTTTTGT 158335
QY 444 AATGTCATTAATTAATAAAAAAAAAAACCAATPACCAACAGACTTACTTCCCTTGGAAA 503
DB 158334 ACTTGCTAGTGTGTTAAGATCACTAATAATCAATGATCTATTTTCCCTTGAAAAG 158275
QY 504 GCACATTTACAAAGG---CTGATCTTAGCCTTTATATTAACAATAAAGAAATACCA 560
DB 158274 GCACAGTTTACAGGAGCCAACTTCCTCTTATTTGTAATAAATAAATACCGA 158215
QY 561 GGTCCCGATATAGCTGTAATTTTATCTTA-----AAAGACAGAAACTTCACTATG 613
DB 158214 AGTCCGACCTATTTGTATATTTTATTCCTAAAGAAAAACAGAACTTCACTGTA 158155
QY 614 CTTTAAATTAAGTATACCTCAGATACCTGCAAGCTTGAAGCTTAACAAAGAGG-- 671
DB 158154 CTTCAACATTAAGTATTAATCACTCAGATATTTTCCAGCTTAGCAGCAAAATCAATT 158095
QY 672 ACAGACAAAGAGCCCA-GCTACTCTTAGAAA---TAAATAGGTGTGCTCTGTA 726
DB 158094 TCAGACAAAGAGATCAACTGCTCTCTAGAGAAATACTTAATGAGGTGAGCTTAAG 158035
QY 727 CATGCCAGGGCTTTGTGCTGTGTG-TGTTCCAGAAAGCTTGCAGTACACAGTC 785
DB 158034 AATGCCAGGGGCTCTGTAAAGATCGTTTTTCCAGAGGGTTCGTGAGCATGGGTC 157975
QY 786 CTTTGGGAGTAAGACATATGCTGATTTTTCGTTGCTGCTAGTGAAGCCCTCAG 845
DB 157974 TGGTTGAGAGGAGGATCTGCTGATTTTTCGTTGCTGCTGCTAGTGAAGCCCTCAG 157915
QY 846 AGAAGATAGTGGTGTGAGCCAG-GGGCGAGAGCGGTGCTGACATGTCGTGCTCT 904
DB 157914 AGAAGATTAAGCGCTTAAGCCAGAGGGGGAGAGCCCACTACATGCTGGCTGCTCT 157855
QY 905 TATCAACTTATCATATTAAGGAAAGAAAGTATGATTGGATCTGACATCTGTAGACT 964
DB 157854 TATCAACTTATCATATTAAGGAAAGAAAGTATGATTGGATCTGACATCTGTAGACT 157795
QY 965 AGGGGAGAAACAAGAGTCGTGAGACCTCTGAGATGAGAGGGGTGCTCTCTCTCTC 1024
DB 157794 TGGGAGAGAGAAACAAGAGCCGTGAAGCTGCTGTAAAGCTGAACACAGCCTCCC 157735
QY 1025 AGGACAGC 1032
DB 157734 AAGTGAGC 157727
```

ABO81802/c	standard; DNA; 240825 BP.
ABO81802;	
14-NOV-2002	(first entry)
Human PG-3 gene SEQ ID NO:1.	
Human; PG-3; chromosome 8; chromosome 8p23; polymorphic; SNP; single nucleotide polymorphism; biallelic marker; DNA repair; recombination; cell cycle control; gene; ds.	
Homo sapiens.	
Key	Location/Qualifiers
5'UTR	1..2000
primer_bind	/*tag= a 1823..1840
primer_bind	/*tag= df 1980..1998
misc_binding	/*tag= dg 1987..2011
misc_feature	/*tag= nh /note= "probe" 1999
primer_bind	/*tag= ad /standard_name= "single nucleotide polymorphism (SNP)" /note= "polymorphic base G or C; the nucleotide is given as an IUPAC ambiguity code in the specification" complement(2000..2018)
exon	/*tag= dh 2001..2079
intron	/*tag= b /label= A 2080..4626
primer_bind	/*tag= c /number= 1 complement(2108..2125)
primer_bind	/*tag= di 4459..4577
primer_bind	/*tag= dj 4582..4600
misc_binding	/*tag= dk 4589..4613
misc_feature	/*tag= ni /note= "probe" 4601
primer_bind	/*tag= ae /standard_name= "single nucleotide polymorphism (SNP)" /note= "polymorphic base A or G; the nucleotide is given as an IUPAC ambiguity code in the specification" complement(4602..4620)
exon	/*tag= di 4627..4718
intron	/*tag= d /label= B 4719..10114
primer_bind	/*tag= e /number= 2 complement(4891..4908)
primer_bind	/*tag= dm 10007..10025
exon	/*tag= dn 10115..10233
primer_bind	/*tag= f /label= C 10209..10227
misc_binding	/*tag= do 10216..10240
	/*tag= nj /note= "probe"

FT	misc_feature	10228	/tag= af	/standard_name= "single nucleotide polymorphism (SNP)"	/note= "polymorphic base G or T; the nucleotide is given as an IUPAC ambiguity code in the specification"
FT	primer_bind		/tag= dp	10229. .10247)	
FT	intron		/tag= dp	10234. .26809	
FT	primer_bind		/number= 3	10267. .10285	
FT	misc_binding		/tag= dq	10274. .10298	
FT			/tag= nk		
FT	misc_feature		/note= "probe"	10286	
FT			/tag= ag		
FT			/standard_name= "single nucleotide polymorphism (SNP)"		
FT			/note= "polymorphic base G or T; the nucleotide is given as an IUPAC ambiguity code in the specification"		
FT	primer_bind		/tag= dr	10370	
FT	misc_feature		/tag= ah		
FT			/standard_name= "single nucleotide polymorphism (SNP)"		
FT			/note= "insertion of G; the nucleotide is given as an IUPAC ambiguity code in the specification"		
FT	primer_bind		/tag= ds	26810. .26897	
FT	exon		/tag= h		
FT			/label= D	26898. .31356	
FT	intron		/tag= i		
FT			/number= 4	31357. .31471	
FT	exon		/tag= j		
FT			/label= E	31472. .34260	
FT	intron		/tag= k		
FT			/number= 5	34261. .34404	
FT	exon		/tag= l		
FT			/label= f	34405. .37376	
FT	intron		/tag= m		
FT			/number= 6	37377. .37466	
FT	exon		/tag= n		
FT			/label= S	37467. .39703	
FT	intron		/tag= o		
FT			/number= 7	39556. .39574	
FT	primer_bind		/tag= dt	39704. .40858	
FT	exon		/tag= p		
FT			/label= T	39877. .39896	
FT	primer_bind		/tag= du	39925. .39943	
FT	primer_bind		/tag= dv	39932. .39956	
FT	misc_binding		/tag= nl		
FT			/note= "probe"	39944	
FT	misc_feature		/tag= ai		
FT			/standard_name= "single nucleotide polymorphism (SNP)"		
FT			/note= "polymorphic base G or T; the nucleotide is given as an IUPAC ambiguity code in the specification"		
FT	primer_bind		/tag= dw	39953. .39970)	
FT	primer_bind				
FT	primer_bind				

```

FT primer_bind /+tag= dx
FT 39954. .39972
FT /+tag= dy
FT 39961. .39965
FT /+tag= nm
FT /note= "probe"
FT 39973
FT /tag= a_j
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base G or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /+tag= dz
FT /complement(40242. .40259)
FT /+tag= ea
FT 40859. .50435
FT /+tag= g
FT /number= 8
FT 41137. .41154
FT /+tag= eb
FT 41366. .41384
FT /+tag= ec
FT 41373. .41397
FT /+tag= nm
FT /note= "probe"
FT 41385. .41403
FT /+tag= ed
FT 41385
FT /tag= ak
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /+tag= ee
FT 41392. .41416
FT /+tag= no
FT /note= "probe"
FT 41404
FT /+tag= a_l
FT /standard_name= "single nucleotide polymorphism (SNP)"
FT /note= "polymorphic base A or C; the nucleotide is given
FT as an IUPAC ambiguity code in the specification"
FT primer_bind /+tag= ef
FT /complement(41564. .41581)
FT /+tag= eg
FT 42122. .42141
FT /+tag= eh
FT 42213. .42231
FT /+tag= ei
FT 42220. .42244
FT /+tag= np
FT /note= "probe"
FT 42232
FT /+tag= am
FT /standard_name= "single nucleotide polymorphism (SNP)"

Query Match 19.5%; Score 238; DB 6; Length 240825;
Best Local Similarity 64.3%; Pred. No. 2,1e-51;
Matches 545; Conservative 0; Mismatches 265; Indels 38; Gaps 11;

QY 221 CCCGAGGAAACCTGTATACAGTAAAGGAGTGTATGTTCTCTCGACATACT 280
DB 158572 CTCAAAAAAGGTTGATTCCTCTTCTACACAGACACTT 158513
QY 281 TCACCTAATATATCTTACTTAACAGACATTAATCGGCCAACACTTAAGTTT 340
DB 158512 AAATTCATATCTTATGTAATGTAATGAACCTTAATCAGTATTTTGAAGT 158453
QY 341 TATTTGTATATGTTTGTGTTAAGGACACAGCATGTGAGTGAAGAACTGATGTTG 400
DB 158452 GAATATCCAGTGAAGTCTTTTAAAG-ATAAAGGTCAAGCTCAG-AAACTGATGTTT 158395

```

```

QY 401 GTAACCTGATTTAATA-----TATCAACTGGGTTAAATTAATA-----AA 443
DB 158394 CTACGCTGACTTAATAACAGATCTTCTATCCAACTGTCGGAATGAATACTTTGTA 158335
QY 444 AATGTCATTAACCTTAATAAAAAAAAAAACCAATACCACAAAGACTTACTCTCCCTGGAAAA 503
DB 158334 ACTTGCTAGTGGTGTCAAGAAATCACTAAATCAAGAAATCTATATTTTCCCTGTGAAG 158275
QY 504 GCACATTTTCAAGG---CTGATCTTAGCCTTATATTAACAATAAGAAATAAACCA 560
DB 158274 GCACAGTTTACAGAGCCAACTTCTCTCTTATTTGTATACAAAAATAAACCA 158215
QY 561 GGTCCGATATAGCTGTAAATTTATCCTA-----AAGAACAGAACTTCACTATG 613
DB 158214 AGTCCGACCTATTTGTATATTTTATCTTAAGAAATAAACAGAACTTCACTATGTA 158155
QY 614 CTTTAAATTAATAGATTAACCTGAGATCTCTGCAAGCTTGGCCACAAAGAGAG-- 671
DB 158154 CTTCAACATTAAGTATTAATTAATTAATTTTGCAGGCTTACGACGCAAAAATAGTT 158095
QY 672 ACAGACAAAGAGCCCA-GCTACTCTAGAGAA-----TAATTAGGCTGTCCTCTGA 726
DB 158094 TCAGACAAAGAGATTAACCTGCTCTCTCTAGAGAAATTAATTTGGGCTGCTCAGGA 158035
QY 727 CATGCCAGGGGCTTTGTGGCTGCTG-TGTTCCAGAGGCTTGTGACATACAGATC 785
DB 158034 AATGCCAGGGGCTCTGTAACAGATCGGTTTTCCAGAGGGTTCTGACGATGGGTC 157975
QY 786 CTTTGGGCGATGAGACATATGCTGTGATTTTTCCTGTCCTGCTAGTACCCTCTAC 845
DB 157974 TGGTTGAGAGGAGAGGATTTCTCTGATTTTTCCTGTCCTGCTAGTACCCTCTAC 157915
QY 846 AGGAAGATAGTGGGTGAGCCAGG-GGGCGAGAGGCTGCTGCACTGTGCTGCTCT 904
DB 157914 AGGAAGATTAACGCTTAAGCCAGAGGCGGACAGCCCACTACACATGCTGCTCTCT 157855
QY 905 TATCAACTTATCATATATAGGAGAAAGAAAGTATGATTTGATCTGACACTGTAGACTC 964
DB 157854 TATCAACTTATCATATATAGGAGAAAGAAAGTATGATTTGATCTGACACTGTAGACTC 157795
QY 965 AGGCGAGAAACAAAGTCCCTGTCACACCTCTGAGTGAAGAGGCTGCTCTCTCTCTC 1024
DB 157794 TGGGAGAGAGGAAACAAAGACCGTGAAGCTGCTCTGTAAAGCTGACACAGCCCTCC 157735
QY 1025 AGGACAGC 1032
DB 157734 AAGTGAGC 157727

RESULT 6
ADQ09378
ID ADQ09378 standard; DNA; 2424 BP.
XX AC
XX ADQ09378;
XX
XX 09-SEP-2004 (first entry)
XX
XX Murine Angiopoietin-2 DNA.
XX
XX Mouse; Angiopoietin-2; gene; de; antisense oligonucleotide;
XX phosphothioate linkage; 2'-O-methoxyethyl sugar moiety;
XX 5-methylcytosine; hyperproliferative disorder; cancer; cyostatic.
XX
XX Mus musculus.
XX
XX
XX Key Location/Qualifiers
XX CDS 211..1701
XX FT /+tag= a
XX FT /product= "Murine Angiopoietin-2"
XX
XX
XX US2004115640-A1.
XX
XX 17-JUN-2004.

```

XX 11-DEC-2002; 2002US-00317803.
PF 11-DEC-2002; 2002US-00317803.
PR 11-DEC-2002; 2002US-00317803.
XX (ISIS-) ISIS PHARM INC.
PA
XX Myers K, Dobie KM;
PI
XX WPI; 2004-449380/42.
DR P-PSDB; ADO09609.
DR GENBANK; AF004326.
XX
PT New oligonucleotide compound that inhibits expression of Angiopoietin-2,
PT useful for preparing a composition for treating hyperproliferative
PT disorder, e.g., cancer.
XX
PS Example 13; SEQ ID NO 11; 102bp; English.
XX
CC The invention relates to a compound targeted to a nucleic acid molecule
CC encoding the human Angiopoietin-2 polypeptide. The compound is an
CC antisense oligonucleotide that specifically hybridizes with the nucleic
CC acid and inhibits expression of the polypeptide. The antisense
CC oligonucleotide comprises at least one modified internucleoside linkage
CC 1.e. a phosphorothioate linkage, at least one modified sugar moiety,
CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified
CC nucleobase comprising a 5-methylcytosine. The antisense compounds are
CC useful for modulating the expression of the human Angiopoietin-2
CC polypeptide and in preparation of a composition for treating
CC hyperproliferative disorders, e.g., cancer. This sequence represents DNA
CC encoding the murine Angiopoietin-2 polypeptide of the invention.
XX
SQ Sequence 2424 BP; 709 A; 581 C; 625 G; 508 T; 0 U; 1 Other;
Query Match 17.5%; Score 213; DB 12; Length 2424;
Best Local Similarity 100.0%; Pred. No. 1.2e-45;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1008 GGCTGCTCTTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGAGC 1067
DB 1 GGCTGCTCTTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGAGC 60
QY 1068 AGGCACCTGGGAAAGAGCTTGTGCGGACGAGAGAGCTTCTCAGTGTGACTTATTGAC 1127
DB 61 AGGCACCTGGGAAAGAGCTTGTGCGGACGAGAGAGCTTCTCAGTGTGACTTATTGAC 120
QY 1128 ACGGCACAGCCTGTGCTTGTAGACAGAGCTGAGAGCTGAGAGCCCACTTGTGAACT 1187
DB 121 ACGGCACAGCCTGTGCTTGTAGACAGAGCTGAGAGCTGAGAGCCCACTTGTGAACT 180
QY 1188 CACAGTTTGAACCCAAAAGAGAGAGAGATG 1220
DB 181 CACAGTTTGAACCCAAAAGAGAGAGAGATG 213
RESULT 7
ADSI3802
ID ADSI3802 standard; DNA; 2424 BP.
XX
AC ADSI3802;
XX
DT 02-DEC-2004 (first entry)
XX
DE Mouse angiotensin-2 (Ang-2) encoding DNA.
XX
XX Ang-1; extracellular matrix; ECM; angiotensin; cytosolic; vasotropic;
KM antidiabetic; antiarthritic; cerebroprotective; antiangiogenic;
XX gene therapy; mouse; gene; Ang-2; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 211..1701

FT /tag= a
FT /product= "Ang-2"
XX
PN MO2004076650-A2.
XX
PD 10-SEP-2004.
XX
PF 27-FEB-2004; 2004WO-US006101.
XX
PR 27-FEB-2003; 2003US-0450582P.
XX
PA (UYBE-) UNIV PENNSYLVANIA.
XX
PI Yu Q;
XX
DR WPI; 2004-653413/63.
XX P-PSDB; ADSI3784.
XX
PT New pharmaceutical composition comprising a pharmaceutical carrier and an
PT amount of an extracellular matrix (ECM)-binding fragment of Ang-1 protein
PT or a mutant Ang-1, useful for treating, e.g., cancer, vascular disease, or
PT ischemia.
XX
PS Disclosure; SEQ ID NO 34; 114bp; English.
XX
CC The invention relates to a pharmaceutical composition comprising a
CC pharmaceutical carrier and an amount of an extracellular matrix (ECM)-
CC binding fragment of angiotensin (Ang)-1 protein or a mutant Ang-1. Also
CC provided are methods of treating an individual suspected of having
CC coronary artery disease, vascular disease or a condition involving
CC ischaemia, of promoting angiogenesis, endothelial survival and
CC maintaining vascular integrity in an individual; of treating an
CC individual suspected of having a disease related to lack of blood vessels
CC to effectively promote angiogenesis in the patients with the diseases
CC related to lack of blood vessels such as ischaemia in hearts and limbs;
CC to reduce stroke, heart attack, blood vessel blockage, haemorrhage,
CC atherosclerosis risk by maintaining the health and integrity of blood
CC vessels; to assist the recovery of the patients who had stroke and the
CC angioplasty procedure by promoting the growth/survival of endothelial
CC cells and establish endothelial monolayer and inhibit excessive
CC inflammation, haemorrhage, and proliferation of vascular smooth muscle;
CC to treat patients with restenosis by inhibiting re-closure of blood
CC vessel after inserting stents into blood vessels; to make stable and
CC functional artificial blood vessels comprising using the composition
CC above; of identifying compounds that modulates binding of Ang-1 to ECM;
CC of treating an individual suspected of having cancer; of preventing
CC diabetes and/or arthritis in an individual suspected of being at risk of
CC developing diabetes or arthritis. The pharmaceutical composition is
CC useful for treating diseases and disorders, e.g., cancer, coronary artery
CC disease, vascular disease, ischaemia, restenosis, diabetes, stroke,
CC angiogenesis, or arthritis. The present sequence represents the mouse Ang
CC -2 encoding DNA.
XX
SQ Sequence 2424 BP; 709 A; 581 C; 625 G; 508 T; 0 U; 1 Other;
Query Match 17.5%; Score 213; DB 13; Length 2424;
Best Local Similarity 100.0%; Pred. No. 1.2e-45;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1008 GGCTGCTCTTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGC 1067
DB 1 GGCTGCTCTTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGC 60
QY 1068 AGGCACCTGGGAAAGAGCTTGTGCGGACGAGAGAGCTTCTCAGTGTGACTTATTGAC 1127
DB 61 AGGCACCTGGGAAAGAGCTTGTGCGGACGAGAGAGCTTCTCAGTGTGACTTATTGAC 120
QY 1128 ACGGCACAGCCTGTGCTTGTAGACAGAGCTGAGAGCTGAGAGCCCACTTGTGAACT 1187
DB 121 ACGGCACAGCCTGTGCTTGTAGACAGAGCTGAGAGCTGAGAGCCCACTTGTGAACT 180
QY 1188 CACAGTTTGAACCCAAAAGAGAGAGAGATG 1220

Db 181 CACAGTTTGAAACCCAAAAGAGAGAGAGATG 213

RESULT 8
ADU39008
ID ADU39008 standard; cDNA; 2424 BP.
XX
AC ADU39008;
XX
DT 27-JAN-2005 (first entry)
XX
DE Mouse Ang2 cDNA.
XX
KW small interfering RNA; siRNA; Ang1; Ang2; Tie2; angiogenesis; inhibitor;
KW type I diabetes; antiangiogenic; cytoprotective; antidiabetic;
KW ophthalmological; nephrologic; neuroprotective; cardiant; vasotropic;
KW gene therapy; cholesterol; endometrial neovascularisation; tumour;
KW diabetic retinopathy; age-related macular degeneration; gene; ss.
XX
OS Mus musculus.
XX
PN WO2004094606-A2.
XX
PD 04-NOV-2004.
XX
PF 19-APR-2004; 2004WO-US012072.
XX
PR 18-APR-2003; 2003US-0463981P.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI Reich SJ, Tolentino MJ;
XX
DR WPI; 2004-795558/78.
XX
PT New isolated small interfering RNA (siRNA) comprising a sense RNA strand
PT and an antisense RNA strand, useful for treating angiogenic disease, e.g.
PT diabetic retinopathy or age-related macular degeneration.
XX
PS Claim 2; SEQ ID NO 5; 182pp; English.

The invention relates to a novel isolated small interfering RNA (siRNA) comprising a sense RNA strand and an antisense RNA strand. The sense and antisense RNA strands form an RNA duplex, where the sense RNA strand comprises a nucleotide sequence substantially identical to a target sequence of 19-25 contiguous nucleotides in human Ang1, Ang2 or Tie2 mRNA, or an alternative splice form, mutant or cognate. The invention further comprises a recombinant plasmid comprising nucleic acid sequences for expressing an siRNA comprising a sense RNA strand and an antisense RNA strand; a recombinant viral vector comprising nucleic acid sequences for expressing an siRNA comprising a sense RNA strand and an antisense RNA strand; a pharmaceutical composition comprising siRNA, the plasmid, or the viral vector and a pharmaceutical carrier; a method of inhibiting expression of human Ang1, Ang2 or Tie2 mRNA, or an alternative splice form, mutant, or cognate; a method of inhibiting angiogenesis in a subject; a method of treating an angiogenic disease in a subject; and a method of treating complications arising from type I diabetes in a subject. The siRNA oligos have the following activities: antiangiogenic, cytoprotective, antidiabetic, ophthalmological, nephrologic, neuroprotective, cardiant, and vasotropic. The siRNA molecules may be used in gene therapy. The siRNA is useful in inhibiting angiogenesis in a subject. The angiogenesis is pathogenic or non-pathogenic. The non-pathogenic angiogenesis is associated with production of fatty tissues or cholesterol production. The non-pathogenic angiogenesis comprises endometrial neovascularisation. The angiogenic disease comprises a tumour associated with a cancer selected from breast cancer, lung cancer, head and neck cancer, brain cancer, abdominal cancer, colon cancer, colorectal cancer, oesophagus cancer, gastrointestinal cancer, glioma, liver cancer, tongue cancer, neuroblastoma, osteosarcoma, ovarian cancer, pancreatic cancer, prostate cancer, retinoblastoma, Wilms tumour, multiple myeloma, skin cancer, lymphoma, or blood cancer. The angiogenic disease may also be diabetic retinopathy or age-related macular degeneration. Treating complications arising from type I diabetes in a subject comprises

CC administering a subject in need of treatment an amount of siRNA
CC comprising a sense RNA strand and an antisense RNA strand. The
CC complications arising from type I diabetes are selected from diabetic
CC retinopathy, diabetic neuropathy, diabetic nephropathy, or macrovascular
CC disease. The macrovascular disease is coronary artery disease.
CC cerebrovascular disease, or peripheral vascular disease. This
CC polynucleotide sequence represents the cDNA of mouse Ang2 of the
CC invention.
XX
SQ Sequence 2424 BP; 709 A; 581 C; 625 G; 508 T; 0 U; 1 Other;
XX

Query Match 17.5%; Score 213; DB 13; Length 2424;
Best Local Similarity 100.0%; Pred. No. 1.2e-45;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1008 GCGTGTCTCTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGAC 1067
Db 1 GCGTGTCTCTCTCTCTCAGACAGCTCCGAGTGTGCGGGAGAGAGAGAGAGAC 60

QY 1068 AGGCACTGGGAAAGAGCCTGTGCGGACGAGAGAGCTCTCAGTATGACTTATTCAC 1127
Db 61 AGGCACTGGGAAAGAGCCTGTGCGGACGAGAGAGCTCTCAGTATGACTTATTCAC 120

QY 1128 ACGGACAGACCTGTGCTTAAACAGACAGCTAGAGACTAGAGCCGAAGTTGCTGAACT 1187
Db 121 ACGGACAGACCTGTGCTTAAACAGACAGCTAGAGACTAGAGCCGAAGTTGCTGAACT 180

QY 1188 CACAGTTTGAAACCCAAAAGAGAGAGATG 1220
Db 181 CACAGTTTGAAACCCAAAAGAGAGAGATG 213

RESULT 9
ADQ09371
ID ADQ09371 standard; DNA; 62705 BP.
XX
AC ADQ09371;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human Angiopoietin-2 DNA #4.
XX
KW Human; Angiopoietin-2; gene; ds; antisense oligonucleotide;
KW phosphorothioate linkage; 2'-O-methoxyethyl sugar moiety;
KW 5-methylcytosine; hyperproliferative disorder; cancer; cytoprotective.
XX
OS Homo sapiens.
XX
PN US2004115640-A1.
XX
PD 17-JUN-2004.
XX
PF 11-DEC-2002; 2002US-00317803.
XX
PR 11-DEC-2002; 2002US-00317803.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Myers K, Dobie KW;
XX
DR WPI; 2004-449380/42.
XX
PT New oligonucleotide compound that inhibits expression of Angiopoietin-2,
PT useful for preparing a composition for treating hyperproliferative
PT disorder, e.g., cancer.
XX
PS Claim 1; SEQ ID NO 4; 102pp; English.

The invention relates to a compound targeted to a nucleic acid molecule encoding the human Angiopoietin-2 polypeptide. The compound is an antisense oligonucleotide that specifically hybridises with the nucleic acid and inhibits expression of the polypeptide. The antisense oligonucleotide comprises at least one modified internucleoside linkage

CC i.e. a phosphorothioate linkage, at least one modified sugar moiety,
 CC preferably a 2'-O-methoxyethyl sugar moiety, or at least one modified
 CC nucleobase comprising a 5-methylcytosine. The antisense compounds are
 CC useful for modulating the expression of the human Angiopoietin-2
 CC polypeptide and in preparation of a composition for treating
 CC hyperproliferative disorders, e.g. cancer. This sequence represents DNA
 CC encoding a human Angiopoietin-2 polypeptide of the invention.

SO Sequence 62705 BP, 19403 A, 12918 C, 12719 G, 17665 T, 0 U, 0 Other;

Query Match 17.4%; Score 211.8; DB 12; Length 62705;
 Best Local Similarity 74.3%; Pred. No. 1e-44;

Matches 335; Conservative 0; Mismatches 107; Indels 9; Gaps 5;

QY 591 AAAGAACAACTTTTACTATGCTTTAAATTAAAGTATTAAGTACTCTGCA 650
 DB 31 AAAAAGACGACCTTCTATGTTACTTCAATTAAGTTATTAAGTATTTGGCA 90
 QY 651 GCTTAGCCTCAAAACGAGCAG--ACAGACAAACAGACCCCA-GCTACTCTTAGGAA-- 705
 DB 91 GCTTAGCAGCGCAAAATTCAGTTTCAGACAAAGATCAACTGCTCTCTAGGAAATA 150
 QY 706 --TAATTAAGGTGTGCTCTGACATGCCAGGGGTCTTGCTGCTGCTG--TGTTCCCA 762
 DB 151 CTAAATGGGGGTGTGCTCTAGGAAATGGCCAGGGGTCTCTAGATGCTTTTCCCA 210
 QY 763 GAAGGCTTCTGACATACAGTCTTTGGGGCAGTAAGCACTATGCTTGAATTTTCTG 822
 DB 211 GAGGGTTTCTGACGATGGGTCTGTTGGAGGGCAGGCAATCTGCTGATTTTTCG 270
 QY 823 TTGCTGCTAGTACACCCCTACAGAGAGTAGTGGTAGGCAAG--GGGCGGACCGGT 881
 DB 271 TTGCTGCTAGTACACCCCTACAGAGAGTAGTGGTAGGCAAG--GGGCGGACCGGT 330
 QY 882 GGCTGACATGCTGTGCTCTCTTATCAACTATATAGGAAAGAAAGTATTGAT 941
 DB 331 CACTACACATGCTGTGCTCTCTTATCAACTATATAGGAAAGAAAGTATTGAT 390
 QY 942 TCGATATCTGACACTGTAGTACTCAGGGGAGAAAGAGTCCGTCAGACCTCTGAGT 1001
 DB 391 TCGATATCTGACACTGTAGTACTCAGGGGAGAGAAAGAGACCGTGAAGCTGCTCT 450
 QY 1002 GAGCAGGGGTGCTCTCTCTCTCAGAGACAGC 1032
 DB 451 GTAAAAGCTGACACAGCCCTCCCAAGTGAGC 481

RESULT 10

ID ABX63024 standard; cDNA, 3251 BP.

AC ABX63024;

DT 25-FEB-2003 (first entry)

DE Human cDNA #24 differentially expressed in activated vascular tissue.

KM Human; gene; ss; vascular tissue; cytosolic; atherosclerosis; cardiac;
 KM hypotensive; antidiabetic; gynaecological; vasotropic; cerebroprotective;
 KM gene therapy; vascular disease; cancer; coronary; artery disease;
 KM hypertension; diabetes; pre-eclampsia; restenosis;
 KM ischaemia-reperfusion injury; stroke.

OS Homo sapiens.

PN US2002137081-A1.

PD 26-SEP-2002.

PF 08-JAN-2002; 2002US-00044090.

PR 28-JUL-2000; 2000US-0222469P.

PR 08-JAN-2001; 2001US-0260483P.

XX (BAND/) BANDMAN O.
 PA Bandman O;
 PI Bandman O;
 DR WPI, 2003-110597/10.

PT Combination for diagnosing, staging, treating, or monitoring the
 PT progression of treatment of a vascular disease, e.g. atherosclerosis,
 PT comprises several cDNAs that are differentially expressed in activated
 PT vascular tissue.

PS Claim 1; Page: 18pp; English.

CC This invention relates to a combination comprising several cDNAs that are
 CC differentially expressed in activated vascular tissue. The invention also
 CC discloses a high throughput method for detecting differentially expressed
 CC cDNAs in a sample. The cDNAs of the invention may have
 CC antiarteriosclerotic, cytosolic, cardiac, hypotensive, antidiabetic,
 CC gynaecological, vasotropic and cerebroprotective activities and may be
 CC used in gene therapy. The cDNAs of the invention may be used in a high-
 CC throughput methods for detecting differential expression of one or more
 CC cDNAs in a sample, or screening several molecules or compounds to
 CC identify a molecule or compound that specifically binds a cDNA of the
 CC invention. A protein encoded by the cDNA may be used to screen several
 CC molecules or compounds to identify a ligand that specifically binds to
 CC the protein, or to produce or purify an antibody to the protein that can
 CC be used to detect a protein in a sample or purify a natural or
 CC recombinant protein from a sample. The nucleotides may be useful for
 CC diagnosing, staging, treating, or monitoring the progression of treatment
 CC of a vascular disease, e.g. atherosclerosis, cancer, coronary artery
 CC disease, hypertension, diabetes, pre-eclampsia, ischaemia-reperfusion
 CC injury, restenosis, or stroke. The cDNAs can also be used for large-scale
 CC genetic or gene expression analysis of several new nucleic acid
 CC molecules. Antibodies to the proteins encoded by the cDNAs are useful for
 CC diagnosing pre-pathologic disorders, and chronic or acute diseases
 CC associated with abnormalities in the expression, amount or distribution
 CC of the protein. The present sequence represents a cDNA of the invention
 CC that is differentially expressed in activated vascular tissue. Note: The
 CC sequence data for this patent did not form part of the specification, but
 CC was obtained in electronic format directly from USPTO at
 CC http://seqdata.uspto.gov/sequence.html?DocID=20020137081

SO Sequence 3251 BP, 1083 A, 645 C, 674 G, 849 T, 0 U, 0 Other;

Query Match 12.8%; Score 156.6; DB 8; Length 3251;
 Best Local Similarity 74.9%; Pred. No. 1.2e-30;

Matches 209; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 755 TGTTCACAGAGGCTTCTGACATACAGTCTTTGGGAGAGTAAAGCACTATGCTTGAT 814
 DB 12 TTTTCCAGAGGGGTTTCTGACAGATGGTCTGTGTGAGGGCAGCACTTCTGCTGAT 71
 QY 815 TTTTCCGTTGCTGCTGCTAGTACCCCTACAGAAAGATAGTGGTAGCCAG--GGGCG 873
 DB 72 TTTTCCGTTGCTGCTGCTAGTACCCCTACAGAAAGATAGTGGTAGCCAGGGGCG 131
 QY 874 GAGCGGCTGCTGACATGCTGCTGCTTATCACTTATATATAGGAAAGAAAG 933
 DB 132 GAGCAGCCACATACATGCTGCTGCTTATCACTTATATATAGGAAAGAAAG 191
 QY 934 TGATTGATTCGATACATGACATGTAAGTACAGGGGAGAAACAAAGAGTCCGTGAGAC 993
 DB 192 TGATTGATTCGATACATGACATGTAAGTACAGGGGAGAGAAACAAAGAGCCTGAAA 251
 QY 994 TCTGAGTACAGAGGCTGCTCTCTCTCAGAGACAGC 1032
 DB 252 GCTGCTCTGTAAAGCTGACACAGCCCTCCAAAGTGAAC 290

RESULT 11

ID ADD69421 standard; DNA, 763 BP.

```
XX AC ADD69421;
XX DT 15-JAN-2004 (first entry)
XX DE Human Ang-2A DNA - SEQ ID 157.
XX KW fusion protein; fibrinogen-like; coiled-like domain;
XX KW angiotensin-related factor; ARF; angiotensin; Ang-1; Ang-2; Ang-3;
XX KW Ang-4; Ang-2X; vunerary; antiinflammatory; vasotropic; necrosis;
XX KW ischaemia; inflammation; wound healing; CCD; FLD; human; ds; Ang-2A.
XX OS Homo sapiens.
XX PN WO2003048185-A2.
XX PD 12-JUN-2003.
XX PF 21-NOV-2002; 2002WO-US037660.
XX PR 30-NOV-2001; 2001US-0334488P.
XX PA (GENV-) GENVEC INC.
XX PI Kessler PD, Kovsed I;
XX DR WPI; 2003-513736/48.
XX PT New fusion protein comprising a fibrinogen-like or coiled-like domain,
XX PT useful for preparing a composition for treating necrosis, ischaemia or
XX PT inflammation, or for promoting wound healing.
XX PS Disclosure; SEQ ID NO 157; 340bp; English.
XX CC The invention relates to a novel fusion protein comprising a fibrinogen-
XX CC like domain (FLD) or coiled-coil domain (CCD). The domain may be
XX CC identical or homologous to that of an angiotensin-related factor (ARF),
XX CC examples of which include Ang (angiotensin)-1, Ang-2, Ang-3, Ang-4 and
XX CC Ang-2X. The molecules of the invention demonstrate vunerary,
XX CC antiinflammatory and vasotropic activities whilst the fusion protein may
XX CC be useful for preparing a composition for treating necrosis, ischaemia or
XX CC inflammation, as well as for promoting wound healing. The current
XX CC sequence is that of the human Ang-2A DNA of the invention.
XX SQ Sequence 763 BP; 202 A; 183 C; 205 G; 173 T; 0 U; 0 Other;
SQ
Query Match 11.8%; Score 143.4; DB 10; Length 763;
Best Local Similarity 76.7%; Pred. No. 26-27;
Matches 188; Conservative 0; Mismatches 56; Indels 1; Gaps 1;
QY 789 TGGGCGAGTAAGCACTATGCTGATTTTCTGTTGCTGCTGCTAGTGAACCCCTACAG 848
DB 5 TTGGAAGGAGGAGGATTCGCTGATTTTCTGTTGCTGCTGCTAGTGAACCCCTACAG 64
QY 849 AAGATGTGGTGAAGCCAGG-GGGCGAGGCGCTGCTGCACTGTCTGCTCTTAT 907
DB 65 AAGATTAACGCTAAAGCAGAGGCGAGGAGCCACTACACTGTCTGCTGCTCTTAT 124
QY 908 CAACCTATCATATTAAGGAAGAAAGATTTGATTTGAGTACTGACACTGTAGACTCAG 967
DB 125 CAACCTATCATATTAAGGAAGAAAGATTTGATTTGAGTACTGACACTGTAGACTCAG 184
QY 968 GGAGAAACAAGAGTCGTCAGACCTCTGAGTGAAGAGGGCTGCTCTTCTCTCAG 1027
DB 185 GGAGAAAGGAACAAGAGACCGTGAAGAGCTCTGTAAAAAGTGAACAGACCCCTCCAG 244
QY 1028 ACAAGC 1032
DB 245 TGAGC 249
```

RESULT 12
AB067157

```
ID AB067157 standard; DNA; 2846 BP.
XX AC AB067157;
XX DT 26-AUG-2002 (first entry)
XX DE Human angiogenesis associated polynucleotide SEQ ID NO 187.
XX KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
XX KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
XX KW macular degeneration; inflammatory bowel disease; Crohn's disease;
XX KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX KW arteriosclerotic; ds.
XX OS Homo sapiens.
XX PN WO200246454-A2.
XX PD 13-JUN-2002.
XX PF 06-DEC-2001; 2001WO-BP014320.
XX PR 06-DEC-2000; 2000DE-01061338.
XX PA (EPIC-) EPIGENOMICS AG.
XX PI Schacht O;
XX DR WPI; 2002-500450/53.
XX PT New nucleic acid fragments from chemically treated angiogenesis-
XX PT associated genes, useful for determining methylation status, e.g. in
XX PT diagnosis or treatment of cancer.
XX PS Claim 1; SEQ ID NO 187; 41bp + Sequence Listing; German.
XX CC The invention relates to a nucleic acid (I) comprising a segment of 18
XX CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
XX CC having sequences (AB066971-AB067178) or their complements. (1), also
XX CC related oligomers, are used to evaluate the methylation status and/or
XX CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
XX CC diagnosis and treatment of eye diseases, proliferative retinopathy,
XX CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
XX CC diabetic retinopathy, macular degeneration caused by neovascularisation,
XX CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
XX CC Crohn's disease. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at http://wipo.int/pub/published\_pct\_sequences
XX SQ Sequence 2846 BP; 871 A; 45 C; 651 G; 1279 T; 0 U; 0 Other;
SQ
Query Match 11.5%; Score 140.6; DB 6; Length 2846;
Best Local Similarity 62.1%; Pred. No. 26-26;
Matches 295; Conservative 0; Mismatches 164; Indels 16; Gaps 4;
QY 522 ATCTAGCTTATATTAATTAAGAAATTAACCAAGTCCGATATAGCTGTAAT 581
DB 1530 ATTTTATTTTATTTATTTAATTAATAAATGAAGTTTATTTATTTGTAATA 1589
QY 582 TTAATTCCT-----AAAGAACGAAGAACTTTCACATAGCTTTAAATTAAGTATTAC 634
DB 1590 TTTTATTTTAAAGAAAAAATGGAATTTTATTTATTTTAAATTAAGTATTTAT 1649
QY 635 CTCAGATCTCTGCAAGCTTAGCTCAACAAGCAGACAGACAAACAGACCCCACTAC 694
DB 1650 TTTAGATATTTTGTATGATTTAGACGTTAAATTAATTTTGAATTAAGATTAATG 1709
QY 695 TCTCT-----AGGAATTAATTTAGGTTGCTCTCTGACATGCCCAGGGCTTGTGC 747
DB 1710 TTTTATTTTGAAGAAATTTAATTTAGGTTGCTGTTTGAAGAAATTTTGAAGG 1769
QY 748 TGGCTG-TGTTCCGAAGAGGCTTCGAGTACACAGTCTTGTGGGCGAGTAAGACTAT 806
```

Db 1770 AGATCGGTTTTTTTAAAGGGTTTTGTAGATGGTTTTGTTGAGGGTAGATATTT 1829
QY 807 GCTCTGATTTTTTCTGTTGCTGGCTAGTAGACCCCCACAGAAAGATAGGGGTGAGCCA 866
Db 1830 GTTTTGATTTTTTTTGTGTTGGTTAGTAGATTTTTTTATAGAAAGATACGGTTAAGTTA 1889
QY 867 GG-GGGGGGAGCGGCTGGCTGCACATGTGGCTGCTCTTATCACTTATCATATAAGG 925
Db 1890 GGAGGGGGGAGTATTATATATATATGTTGGTTGTTTTTATTATTTATATATAAGGA 1949
QY 926 AAGGAAGTATGATTGCGATCTGACACTGTAGACTCAGGGAGAAACAAAGA 980
Db 1950 AAGGAAGTATGATTGCGATTTGATTTGATTTGAGATTTGGGAGAGAGAAATA 2004

RESULT 13
ABQ67158/c
ID ABQ67158 standard; DNA; 2846 BP.
XX
AC ABQ67158;
XX
DT 28-AUG-2002 (first entry)
XX
DE Human angiogenesis associated polynucleotide SEQ ID NO 188.
XX
KM Human: angiogenesis; methylation; eye disease; glaucoma; tumour;
KM inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
KM macular degeneration; inflammatory bowel disease; Crohn's disease;
KM antithrombotic; antidiabetic; antidiabetic; antidiabetic;
KM antiarteriosclerotic; ds.
XX
OS Homo sapiens.
XX
PN WO20024654-A2.
XX
PD 13-JUN-2002.
XX
PF 06-DEC-2001; 2001WO-EP014320.
XX
PR 06-DEC-2000; 2000DE-01061338.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Schacht O;
XX
DR WPI: 2002-500450/53.
XX
PT New nucleic acid fragments from chemically treated angiogenesis-
PT associated genes, useful for determining methylation status, e.g. in
PT diagnosis or treatment of cancer.
XX
PS Claim 1; SEQ ID NO 188; 41pp + Sequence Listing; German.
XX
CC The invention relates to a nucleic acid (I) comprising a segment of 18
CC bases of chemically pre-treated DNA of angiogenesis-associated genes (II)
CC having sequences (ABQ66971-ABQ67178) or their complements. (I) also
CC related oligomers, are used to evaluate the methylation status and/or
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 2846 BP; 759 A; 45 C; 565 G; 1477 T; 0 U; 0 Other;

Query Match 11.4%; Score 139.4; DB 6; Length 2846;
Best Local Similarity 56.4%; Pred. No. 4,2e-26;
Matches 457; Conservative 0; Mismatches 316; Indels 38; Gaps 9;
QY 206 CCAAGAGTGTATGAGCCGAGGAAACCTGTATACATGTAAGAAAGAAAGGTATGTTTGT 265

Db 1651 CGACAAAACGAAACCTCCGCTCCAAAAAATTTATATATTCCTCTT 1592
QY 266 TCCCTCGACATATCTTCACTCAATATATCTATGTTACTTAAACAGACATATATAGGCCA 325
Db 1591 TCTACACAAACCTTAAATTTCTAATTTCTATATATATATATTAATTAATTAATTAATTA 1532
QY 326 AACCACTTAACTTATATTTGATATGATATTTTGTGTTAAGGACAGATGTGATGCTG 385
Db 1531 ATATATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1474
QY 386 AGAAACTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 435
Db 1473 CAAAAACATATTTTCTAAGCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1414
QY 436 TAAAA-----AAAATGTCATTACTTAAAAAATTAATTAATTAATTAATTAATTAATTA 488
Db 1413 TAAAAATCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1354
QY 489 CTTCCTTGGAAAAAGACA---TTTACAGAGGCTGATCTTAGCTTTATATTTACATA 545
Db 1353 TTTCCTTAAATAAACAACAGCTTTACAAAAACCAAACTTCTCTTTATTTATATA 1294
QY 546 AAGAAATTAACCAAGTCCGATATAGCTGATATTTATTTCT-----AAAAGACA 598
Db 1293 ACAAATTAACCGAAATCTTAACCTTATATATTTTATTTATTTCTTAAAAAATTAACA 1234
QY 599 GAACTTTCACTATGCTTTTAAATTAATGATTAATCTGAGATATCTGCAAGCTTAGC- 657
Db 1233 AAACTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1174
QY 658 CTACAAACGAGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 711
Db 1173 CGACAAAATTAATTTGAAACCAAAAAATCACTACTCTCTTAAATAATTAATTAATTA 1114
QY 712 GGGTGTGCTCTGACATATCCCAAGGGCTTGTGCTGTGTG-TTCCAGAAAGCTT 770
Db 1113 AAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1054
QY 771 CTGCAATTAACAGTCTTTGGGGCAGTAAGCAGTATGATTTTCTGTGCTGTG 830
Db 1053 CTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 994
QY 831 CTAGTGAACCCCTACAGAAAGATAGTGGTGAAGC-AGGGGGGAGAGGCTGGCTGCAC 889
Db 993 CTAAATACCCCTACAAAAAATTAAGACCTAAACCAAAAAAAGAAACCAACCACTACAC 934
QY 890 ATGTCTGCTGCTCTTATCACTTATCACTTATCACTTATCACTTATCACTTATCACTTATCA 949
Db 933 ATATCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 874
QY 950 TGACACTGTAGACTCAGGGAGAAACAAAGA 980
Db 873 TAACTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 843

RESULT 14
ADA71938/c
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KM Plant; bacterial infection; fungal infection; viral infection; rice;
KM gene; ds.
XX
OS Oryza sativa.
XX
PN WO200300898-A1.
XX

```

PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001WO-IB001105.
XX
PR 22-JUN-2001; 2001WO-IB001105.
XX
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
PI Katsigiri F, Qian S, Tao Y, Whitlam S, Xie Z, Zhu T, Zou G;
XX WPI; 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 27; SEQ ID NO 5263; 899pp; English.
XX
CC The present invention relates to a method (M1) for identifying genes
CC involved in plant resistance or response to pathogenic infection. M1
CC comprises identifying a gene whose expression is significantly altered in
CC the incompatible interaction of plant gene expression relative to
CC expression of the gene in an uninfected plant, in a mutant plant that
CC does not express a gene associated with response to pathogenic infection,
CC or in a corresponding incompatible or compatible interaction. (M1) is
CC useful for conferring resistance to resistance or tolerance to a plant to
CC bacterial, fungal or viral infection. The present sequence was used to
CC illustrate the invention.
XX
SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;
XX
Query Match 4.8%; Score 58; DB 8; Length 2000;
Best Local Similarity 11.6%; Pred. No. 0.00013;
Matches 49; Conservative 204; Mismatches 169; Indels 2; Gaps 1;
XX
QY 329 CACCTTAAGTTTATTTAGTATGATTTGTTGTTAAGCAGACAGCTGAGTCTGAGA 388
DB 1009 SWSWYVAAWKKYTKWTATASSTWKWYWAYKRAYASRSRKTWCTGGRMATYCGTKRW 950
QY 389 AAGAGTATGTGTGTAATCTGATTTAATAATCAACCTGGGTTAAATAAAAAAATGT 448
DB 949 AAGRRWRWAMCWYCCMMKMKWKTSCMMWKYRTSCWTTMMGAWRIAYVAMRRRTY 890
QY 449 GCATACTTAAAAAACAACCAATACCAACAAGACTTACTTCCCTTGAAGAACACA 508
DB 889 KMSWRMTWMTKMAWTMTTCOMAKWVAITGMATMMWRMYTYCYVMTCAKCKYKMA 830
QY 509 TTTACAAGGCTGATCTTAGCCTTATATTTACAATAAAGAAAATAAACCAAGTCCGA 568
DB 829 MTKWMTTACAWRAISWRRAAGMKRKYKMKRAVAMWMBRCWKGARMSRWRK 770
QY 569 TATAGCTGTAATTTATTCCTTAAAGAACAGAACTTTACATAGCTTTAAATTAAGT 628
DB 769 KXIARIRYKMAAMTMMWSMRKSYRMWSGMRMWSAWRYCSRKCAKTYVASARAT 710
QY 629 GATTACCTGATCTGCAAGCTTAGCTCAACAAC--GAGCAGACAGACAGAGCC 686
DB 709 KRAKRSYRBRWYKWKWMTYRIRYWRSCMTBRAMSKRRKMAAGSMSCMMYRGAAS 650
QY 687 CCAAGTACTCTTAGAATAATTAAGGCTGCTCTGACATGCCCAAGGGCTTGTGG 746
DB 649 MWSYKYSOSAKCKRTYMTSYMTGYMWSYYSKMSWTSKMSYMGHMTCTMYTSMG 590
QY 747 CTGG 750
DB 589 STR 586

```

RESULT 15
ABL32893/c
ID ABL32893 standard; DNA; 10710 BP.

```

XX ABL32893;
XX
XX 26-MAR-2002 (first entry)
XX
XX Human immune system associated gene SEQ ID NO: 866.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antianaemic; cyostatic; noctropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antineumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful for
XX diagnosis and treatment of diseases associated with abnormal cytosine
XX methylation.
XX
XX Claim 1; SEQ ID NO 866; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention
XX
XX Sequence 10710 BP; 3653 A; 81 C; 1963 G; 5013 T; 0 U; 0 Other;
XX
Query Match 4.5%; Score 55; DB 6; Length 10710;
Best Local Similarity 48.8%; Pred. No. 0.0018;
Matches 148; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
XX
QY 256 TATGTTGTTCTTCGACATATCTTCACTCAATATCTATTGTTACTTAACAGCAATT 315
DB 5616 TATAATTTCTAATCTCACAAATATCCTTATTAATAAAAAAATAATCAAAACAAACCT 5557
QY 316 AATCGGCAAAACCACTTAAGTTTATTTGTAATGTAATTTGTTAAGCAGACAT 375
DB 5556 AATATTTTAATTCCTTTTCAATACATTTATGATTTTATATATATATCTTACATATT 5497
QY 376 GTGAGTGTGAGAAATAGATGTTGTAACTGATTTATATATATCAAACTGGGTTAAA 435
DB 5496 TTAATTAATATAATAATAAAAAATTTTAATAATAATATATATATATATCTATTTATA 5437
QY 436 TAAAAAATAATGTGATTAATTAATAAAAAAACAATATCAACAAGACTTACTTCCC 495
DB 5436 AATATCTTCTTCAACATACATTAACAAAAAATAAATAAATAATTAATCTCTCCATCCG 5377
QY 496 TTGAAAAAGACATTTAACAGGCTGATCTTATGCTTATATATTTCAATTAAGAATA 555
DB 5376 TAAATTAACAAAATTAATTAACCAACAAATAAAAAATATAAAAAATTAATTAATA 5317

```

QY	556	ACC	558
Db	5316	AAC	5314

Search completed: December 12, 2005, 12:16:44
Job time : 762 secs

PUBMED
COMMENT

12466851

Contact: Yoshitake Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Komodo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submision

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES
source

Location/Qualifiers
1..699
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="5031400E18"
/sex="female"
/tissue_type="ovary and uterus"
/dev_stage="11 days pregnant, adult"
/lab_host="DH10B"
/clone_1ib="RIKEN full-length enriched, 11 days pregnant adult female ovary and uterus"
/note="Site 1: SalI, Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTTGATTTATTTATATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI."

ORIGIN

Query Match 22.6%; Score 276; DB 5; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.7e-58;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
945 GATCTGACACTGTAGACTCAGGAGGAGAAAGAGAGCTCGTGAAGTGGAGTGAAG 1004

Db
1 GATCTGACACTGTAGACTCAGGAGGAGAAAGAGAGCTCGTGAAGTGGAGTGAAG 60
QY
1005 CAGGGCTGCTCTTCTCTCTCAGACAGCTCCAGTGTCTCCGGGAGAGAGAGAGA 1064
61 CAGGGCTGCTCTTCTCTCTCAGACAGCTCCAGTGTCTCCGGGAGAGAGAGAGA 120
Db
1065 GACAGGCACTGGGAAAGAGCTGCTCCGAGACGAGAGAGCTCTCAGTGAATG 1124
121 GACAGGCACTGGGAAAGAGCTGCTCCGAGACGAGAGAGCTCTCAGTGAATG 180
QY
1125 CACACGGCAGACAGCCCTGCTGCTTAGACAGCAGCTGAGAGCGCAAGTTGCTGA 1184
181 CACACGGCAGACAGCCCTGCTGCTTAGACAGCAGCTGAGAGCGCAAGTTGCTGA 240
QY
1185 ACTCAGCTTGAAGAACCCAAAAGAGAGAGAAATG 1220
241 ACTCAGCTTGAAGAACCCAAAAGAGAGAGAAATG 276

RESULT 2
AK019860
LOCUS
DEFINITION
AK019860
2475 bp mRNA linear HTC 03-APR-2004
Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
RIKEN full-length enriched library, clone:5031400E18
product:ANGIOPOIETIN-2 PRECURSOR, full insert sequence.
ACCESSION
AK019860.1 GI:12860231
VERSION
AK019860
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Carninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
AUTHORS

TITLE
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159

TITLE
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
PUBMED
11078661
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishii, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Onoda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

TITLE
JOURNAL
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
FUNCTIONAL annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
5

TITLE
JOURNAL
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2475)
JOURNAL
Nature 420, 563-573 (2002)
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,

REFERENCE	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	ORGANISM	Mus musculus
AUTHORS	1 (bases 1 to 354)	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
	Maira, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Stepcoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.	AUTHORS	1 Carninci, P. and Hayashizaki, Y.
TITLE	The WashU-HMI Mouse EST Project	TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Unpublished (1996)	JOURNAL	Methods. Enzymol. 303, 19-44 (1999)
COMMENT	Contact: Maira M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:344811 Seq primer: -28ml3 rev1 ET from Amerisham High quality sequence stop: 255. Location/Qualifiers	AUTHORS	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
FEATURES	1. 354 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="IMAGE:570193" /cissue_type="embryo" /dev_stage="8.5dpc embryo" /lab_host="DH108" /clone_lib="Life Tech mouse embryo 8.5dpc 10664019" /note="Organ: whole embryo; Vector: PCMV-SPORT2; Site: 1: Salt; Site 2: NotI; Cloned unidirectionally. Primer: oligo dt. 8.5dpc embryos. PCMV-SPORT2 vector."	TITLE	Normalisation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
ORIGIN	Query Match 20.7%; Score 253; DB 1; Length 354; Best Local Similarity 100.0%; Pred. No. 9.7e-53; Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	PUBMED	Genome Res. 10 (10), 1617-1630 (2000)
	968 GGAGAAACAAAGAGTCCGTGACAGCTCTGAGTGCAGAGGCTGCTCTTCTTCAGG 1027	JOURNAL	3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura, S., Kawai, D., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
	1 GGAGAAACAAAGAGTCCGTGACAGCTCTGAGTGCAGAGGCTGCTCTTCTTCAGG 60	AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
	1028 ACAGCTCCGAGTGTGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTG 1087	PUBMED	Functional annotation of a full-length mouse cDNA collection
	61 ACAGCTCCGAGTGTGCGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCTG 120	AUTHORS	Nature 409, 685-690 (2001)
	1088 CTGCGGGAGCGAGAGAGCTCTCACTGATGAGCTTATTCACAGCGCAGCCCTGTGCTT 1147	TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
	121 CTGCGGGAGCGAGAGAGCTCTCACTGATGAGCTTATTCACAGCGCAGCCCTGTGCTT 180	JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
	1148 AGACACGAGCTGAGAGCTCGAGAGCGCAAGTTTGTGAACCTCAAGTTTAAACCCAAAA 1207	AUTHORS	6 (bases 1 to 2443)
	181 AGACACGAGCTGAGAGCTCGAGAGCGCAAGTTTGTGAACCTCAAGTTTAAACCCAAAA 240	TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
	1208 GAGAGAGAGATG 1220	JOURNAL	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
	241 GAGAGAGAGATG 253	COMMENT	Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers
RESULT 6	AK048622 2443 bp mRNA linear HTC 03-Apr-2004	FEATURES	1. 2443 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J"
LOCUS	AK048622		
DEFINITION	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130089A05 product: ANGIOPOIETIN-2 PRECURSOR, full insert sequence.		
ACCESSION	AK048622		
VERSION	AK048622.1 GI:26339449		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		

|||||
Db 121 CCTGTGCTTAGACGACGACTGAGAGCCGAACTTCTGAATCTCAGCTTTA 180
QY 1197 GAAACCAAAAAGAGAGAGAAATG 1220
Db 181 GAAACCAAAAAGAGAGAGAAATG 204
RESULT 11
AZ984331 661 bp DNA linear GSS 27-APR-2001
AZ984331
LOCUS 2M0265M18R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION clone UUGC2M0265M18 R, genomic survey sequence.
ACCESSION AZ984331
VERSION AZ984331.1 GI:13855558
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridea; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
Plasmid inserts
Unpublished (2000)
CONTACT: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0265 row: M column: 18
Seq primer: CACACAGAAACAGCTATGACC
Class: Plasmid ends
High quality sequence stop: 661.
Location/Qualifiers
1. .661
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0265M18"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/note="Vector: PWD42nv. Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid p1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

Query Match 16.7%; Score 203.4; DB 9; Length 661;
Best Local Similarity 99.1%; Pred. No. 3.9e-40;
Matches 215; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1004 GCAGGCTGCTCTCTCTCTCAGAGACGCTCCGAGTGTCCCGGAGAGAGAGAG 1063
Db 8 GCCGGCTCTCTCTCTCTCTCA-GACAGCTCCGAGTGTCCCGGAGAGAGAGAG 66
QY 1064 AGACAGGCACTGGAGAAAGACCTGTGCGGAGCGAGAAAGCTCTCAGTGAGCTTAT 1123
Db 67 AGACAGGCACTGGAGAAAGACCTGTGCGGAGCGAGAGAGCTCTCAGTGAGCTTAT 126
QY 1124 TCACAGGCAAGCCCTGTGCTCTTACACAGAGCTCAGAGCTCAGAGACCAAGTTGCTG 1183
Db 127 TCACAGGCAAGCCCTGTGCTCTTACACAGAGCTCAGAGCTCAGAGACCAAGTTGCTG 186
QY 1184 AACTCAGCTTGAACCCAAAAGAGAGAGAAATG 1220
Db 187 AACTCAGCTTGAACCCAAAAGAGAGAGAAATG 223

RESULT 12
CR905089 793 bp DNA linear GSS 23-NOV-2004
LOCUS CR905089
DEFINITION Sus scrofa BES, genomic survey sequence.
ACCESSION CR905089
VERSION CR905089.1 GI:56229586
KEYWORDS GSS; Bac-end sequence BES; Genome Survey Sequence.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
Sus.

REFERENCE
AUTHORS Rogel-Galliard,C., Bourgeaux,N., Billaule,A., Vaiman,M. and
Chardon,P.
TITLE Construction of a swine BAC library: application to the
characterization and mapping of porcine type C endoviral elements
JOURNAL Cytogenet. Cell Genet. 85 (3-4), 205-211 (1999)
PUBMED 10449899
REFERENCE
AUTHORS Chardon,P., Iannucci,I., N., Roig,A., Dosat,C., Demars,J.,
Rogel-Galliard,C., Roy,A., Schibler,L. and Milan,D.
TITLE A physical map of the swine genome
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 793)
GENOSCOPE.
TITLE Direct Submission
SUBMITTED (18-NOV-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Location/Qualifiers
1. .793
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="Large White"
/db_xref="taxon:9823"
/clone="B10352F04"
/sex="male"
/cell_type="fibroblast"
/clone_lib="SBAB"
/note="Genoscope sequence ID : IH0AAJ4DD10FM1"

ORIGIN
source
1. .793
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="Large White"
/db_xref="taxon:9823"
/clone="B10352F04"
/sex="male"
/cell_type="fibroblast"
/clone_lib="SBAB"
/note="Genoscope sequence ID : IH0AAJ4DD10FM1"

Query Match 15.2%; Score 185.8; DB 11; Length 793;
Best Local Similarity 65.5%; Pred. No. 1.1e-35;
Matches 476; Conservative 0; Mismatches 202; Indels 49; Gaps 12;
QY 290 TATGATTTGTTACTTAACAGACATTAATCAGGCCAACCCTTAAGTTTATTTGAT 349
Db 51 TATGATTTAATTAATTAAGAGGCCATTAATCAGTCAGCAGATTTTAATTTACTTGAAT 110
QY 350 A-----GTATTTGTGTGAAGGCACAGACATGTGATGTCGAGAAAACTATTTGGT 402

```

Db      111 ACTGAAGTGAATTTTGTGTTAAAGATTAATGTACAGCTCAAAACCTGATTTTCCT 170
Qy      403 AACTGATTAAATA-----TATCAACTGGGTTAAATAAAAAA----- 445
Db      171 AACCCATTTAATAAGTAGATGTTTCTATCCAAATGTTTAAAGTAGAAAATTCGCAT 230
Qy      446 ---TGTCATTACTTAAAAAACCMAATACCAACAAGCTTACTTCCCTGGAAA 502
Db      231 TCTTAAGCTAGTGTTCAAAAATCAAGTAATACCAAGAACTATTTCCCTATGAAAA 290
Qy      503 AGCACA-----TTTACAAGGCTGATCTTAGCCCTTATATTATTAATAAAGAAATAAAC 558
Db      291 GGCAGAAAGATTATTAAGACAGACACTTCTTCTTTTGTGTAATACAAATAAAC 350
Qy      559 AAGTCCCGATATAGCTGTAATTTATCTT-AAAAGAACAGAACTTCACTAGCTTT 617
Db      351 AAGTCTCTACATATTGTATGTAATTTTATTCCTMAAGAAAAACAGAACCTTCATTG 410
Qy      618 AAAATTAAGTATTAATCTCACTAGTCTGCAAGCTTAAGC--TACAAACGACAGACA 674
Db      411 TACTTAAAGTATGCCCCAGATTTTACAGACTCAGCCCTGTGAAATCAGTTTCAGA 470
Qy      675 GACAAAGAGCCCAAGCTACTCTCTAGAAAT-----ATTAGGGTGGCTCTGACATG 730
Db      471 CACAAAGACCCCAAGCTCTCTAGAAATCTCAATTAAGGAGTGTCTCTGAAATA 530
Qy      731 CCCAGGGTCTTGTGGCTGCTG- TGTTCCAGAAAGGCTTGTGACAGACAGTCT-T 788
Db      531 CCCAGGGGCTGTGAATCATGATGGGCTTTTCCAGAGGGCTTCTACAGCTATAGGCCCTG 590
Qy      789 TGGGACAGTAACGACTATGCTGATTTTCTGTGCTGCTGCTAGTACCCCCCTACAG 848
Db      591 CTGGCAACACACACACACCTGGAATTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 650
Qy      849 AAGATAGTGGTGTGAGCAG-----GGGCGAGAGCGGCTGCTGACATGTCTGCTGT 902
Db      651 AAGATAGGAGCAAGCGGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 709
Qy      903 CTTATCAACTTATTCATTAAGGAAAGAAAGTATGATTCGATTAAGTCTGACCTGTAAC 962
Db      710 CTTATCAACTTATTCATTAAGTATTAAGTATTAAGTATTAAGTATTAAGTATTAAGT 768
Qy      963 TCAGGGG 969
Db      769 TCTGGGG 775

```

```

RESULT 13
BX506958          726 bp      mRNA      linear      EST 04-SEP-2003
LOCUS            DXFZP779M112.F1 779 (synonym: hnccl) Homo sapiens cDNA clone
DEFINITION      DXFZP779M112 5', mRNA sequence.
ACCESSION      BX506958
VERSION        BX506958.1 GI:32047724
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 726)
Pousetka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
EST (Pousetka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
Unpublished (2003)
TITLE          JOURNAL
COMMENT        MIPS

```

Insolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DXFZP779M112) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
location/Qualifiers

FEATURES

source

```

1..726
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DXFZP779M112"
/tissue_type="liver"
/dev stage="fetal"
/lab host="DH10B"
/clone_id="779 (synonym: hnccl)"
/notes="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiIB"

```

ORIGIN

Query Match

Best Local Similarity 75.4%; Pred. No. 7,4e-34;
Matches 248; Conservative 0; Mismatches 79; Indels 2; Gaps 2;

```

Qy      706 TAATTAAGGTGGTGGCTCTGACATGCCCCAGGGGCTTGTGCTGCTGTG-TTCCAGA 764
Db      46 TAATGGGGGTGGTGGCTCTGAGAAATGCCCCAGGGGCTTGTGACATGCGTTTCCAGA 105
Qy      765 AGGCTTGGCAGTACACAGTCTTTGGGCGAGTAAGCACTATGCTCTGATTTTCTGTT 824
Db      106 GGGTTCTGACAGCATGGGCTCTGTTGGAGGGCAGGCAATTTGCTGATTTTTCCTGTT 165
Qy      825 GCTGGCTAGTAGACCCCTTACAGAAAGATAGTGGGTGAGCCAGG-GGGCGAGCGGCTGG 883
Db      166 GCTGGCTAGTAGACCCCTTACAGAAAGATAGCGCTAAGCCAGAGGGCGGAGCAGCCCA 225
Qy      884 CTGACATGCTGGCTGCTCTTATCAACTTATTCATTAAGGAAAGAAAGTATGATTC 943
Db      226 CTACACATGCTGGGTGCTCTTATCAACTTATTCATTAAGGAAAGAAAGTATGATTC 285
Qy      944 GGATTAAGTACATGTAAGCTAGGAGGAGAAACAAGAGTCCGTGCACACCTCTGAGTGA 1003
Db      286 GGATTAAGTACATGTAAGCTAGGAGGAGGAGAAACAAGAGTCCGTGCACACCTCTGT 345
Qy      1004 GCAGGGCTGCTCTTCTCTCTCAGACAGC 1032
Db      346 AAAAGCTGACACAGCCCTCCCAAGTGAGC 374

```

```

RESULT 14
BB601394          204 bp      mRNA      linear      EST 01-DEC-2000
LOCUS            BB601394 RIKEN full-length enriched, 13 days embryo lung Mus
DEFINITION      BB601394 RIKEN full-length enriched, 13 days embryo lung Mus
ACCESSION      BB601394
VERSION        BB601394.1 GI:11509995
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Mus.
1 (bases 1 to 204)
Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T.,
Carninci,P., Hanagaki,T., Hayatsu,N., Hirooka,T., Hirozane,T.,
Hodoyama,Y., Imotani,K., Ishii,Y., Itoh,M., Iwata,M., Kawai,J.,
Kojima,Y., Kono,H., Kusakabe,M., Matsuyama,T., Miyazaki,A.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Okazaki,Y.,
Okido,T., Owa,C., Sakai,C., Sakai,K., Sasaki,D., Sato,K.,
Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagawa,A., Takehashi,F., Tanaka,T., Toya,T.,
Watabiki,A., Yamamura,T., Yasunishi,A., Yoshida,K., Yoshiki,A.,

```


TITLE
JOURNAL
COMMENT

Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Aizawa, K. et al. 2000)
Unpublished (2000)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.jp, URL: http://genome-gsc.riken.jp/
Carinci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Thermolabile and thermolabile activation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Kitsuai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
Tomaru, Y., Carinci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
Okazaki, Y. and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for
further details.

FEATURES
source

Location/Qualifiers
1..204
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D430011L06"
/tissue_type="lung"
/dev_stage="13 days embryo"
/lab_host="DH10B"
/clone_1ib="RIKEN full-length enriched, 13 days embryo
lung"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse cDNAs. 1st strand cDNA was
primed with a primer 15'.
GAGAGAGAGAGCGCGCACTGAGTTTCTTTTCTTTT 3'. cDNA was
transcribed by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence 15'.
GAGAGAGAGATTCTCGAGTTAATTAATTAATTCCTCCCTCCCTCC 3'. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBlueScript KS(+) after bulk excision from Lambda PUC I."

ORIGIN

Query Match 14.4%; Score 175.8; DB 2; Length 204;
Best Local Similarity 91.6%; Pred. No. 2.9e-33;
Matches 186; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
948 ACTGACACTGTAGCTCAGGAGAAAGAGTCCGTCAGACCTTGAGTGAGC 1007
|||||
2 ACTGACACTGTAAATCAGGAGAAAGAGTCCGTCAGACCTTGAGTGAGC 61
|||||
1008 GGCCTCTCTTCTCTCAGAGACGCTCCGAGTGTGCGCGGAGAGAGAGAGAC 1067
|||||
62 GGGTGTCTCTCTCTCAGAGAGCTCCGAGTGTGCGCGGAGAGAGAGAGAGAC 121
|||||
1068 AGGCACTGGGAAGAGCTCTGCGGAGAGAGAGCTCTCAGTATGATCTTATTC 1127
|||||
122 AGGCACTGGGAAGAGCTCTGCGGAGAGAGAGAGCTCTCAGTATGATCTTATTC 181
|||||
1128 ACGGCAAGACCTGTGCTTAGA 1150
|||||
DB 182 ACGGCAAGACCTGTGCTTAGA 204
|||||

RESULT 15
BUS80281 525 bp mRNA linear EST 17-SEP-2002
LOCUS BUS80281
DEFINITION BK79709.y1 Kaestner ngn3 wt Mus musculus cDNA 5' similar to
IR:035608 035608 ANGIOPOIETIN 2.; mRNA sequence.
ACCESSION BUS80281
VERSION BUS80281.1 GI:23065508
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

Melton, D., Brown, J., Kently, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scaer, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistein, A.,
Schmitt, A., Theising, B., Ritzer, E., Ronko, I., Bennett, J.,
Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R.,
Williams, T., Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Unpublished (2000)

JOURNAL

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu

Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
2000) Library was constructed by Catherine Lee DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Marie Scaer
(mscaer@mail.med.upenn.edu)
Seq primer: -40RP from Gibco
(mscaer@mail.med.upenn.edu)
High quality sequence stop: 343.
Location/Qualifiers
1..525

FEATURES
source

/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x CD1"
/db_xref="taxon:10090"
/dev_stage="P.c. 14.5"
/lab_host="E. coli-DH2S (GIBCO)"
/clone_1ib="Kaestner ngn3 wt"
/note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1:
Not I; Site_2: Sal I; The library was prepared by
Catherine S. Lee and has not been published. The pancreas
was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
2000). The cDNA's were prepared with an oligo containing a
NotI site, and SalI linkers were added to the ends. The
inserts were cut with NotI before being cloned into the
NotI-SalI sites in the vectors. This is one of two
libraries, ngn3 wt and ngn3 -/- . The wt library is in
pSPORT1, T7 promoter is 5'."

ORIGIN

Query Match 13.8%; Score 167.8; DB 5; Length 525;
Best Local Similarity 96.1%; Pred. No. 3.7e-31;
Matches 172; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
1042 GCGGGGAG 1101
|||||
35 GCGAGTACGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 94
|||||
1102 AGGCTCTACGATGATGATTTATTCACAGGAGAGAGAGAGAGAGAGAGAGAG 1161
|||||
DB 95 AGGCTCTACGATGATGATTTATTCACAGGAGAGAGAGAGAGAGAGAGAG 154
|||||
QY 1162 AGCTGAGAGCGCAAGTTTGTGCACTCAAGTTTAGAACCCAAAGAGAGAGAGATG 1220
|||||

Db 155 AGCTCAGGAGCGAGTTGCTGTAATCAGTTTGAACCCAAAAGAGAGAGATG 213

Search completed: December 12, 2005, 15:19:28
Job time : 4730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2005, 12:00:02 ; Search time 260 Seconds

(without alignments)

8340.866 Million cell updates/sec

Title: US-10-627-075-1

Perfect score: 1220

Sequence: 1 gagcttccagctaccgac.....ccaaaagagagagagatg 1220

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/1_COMB.seq:*\n2: /cgn2_6/prodata/1/ina/5_COMB.seq:*\n3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*\n4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*\n5: /cgn2_6/prodata/1/ina/H_COMB.seq:*\n6: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*\n7: /cgn2_6/prodata/1/ina/PP_COMB.seq:*\n8: /cgn2_6/prodata/1/ina/RE_COMB.seq:*\n9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	47.4	3.9	7218	2	US-08-232-463-14 Sequence 14, Appl
C 2	46.4	3.8	1864	3	US-09-468-265-4 Sequence 4, Appl
C 3	46.2	3.8	1141	3	US-09-806-708B-22 Sequence 22, Appl
C 4	45.4	3.7	20674	3	US-09-641-638-651 Sequence 651, App
C 5	45.4	3.7	20674	3	US-10-170-097-651 Sequence 651, App
C 6	44.6	3.7	732	3	US-08-998-416-1036 Sequence 1036, App
C 7	44	3.6	1141	3	US-09-806-708B-22 Sequence 22, Appl
C 8	43.8	3.6	2308	3	US-10-104-047-682 Sequence 682, App
C 9	42.6	3.5	80411	3	US-09-949-016-15777 Sequence 15777, A
C 10	42.4	3.5	34393	3	US-09-949-016-12063 Sequence 12063, A
C 11	42.4	3.5	95020	3	US-09-949-016-12372 Sequence 12372, A
C 12	42.2	3.5	50000	3	US-09-662-254B-25 Sequence 25, Appl
C 13	42	3.4	601	3	US-09-949-016-83375 Sequence 83375, A
C 14	42	3.4	601	3	US-09-949-016-83376 Sequence 83376, A
C 15	42	3.4	143550	3	US-09-949-016-14143 Sequence 14143, A
C 16	41.8	3.4	832	3	US-09-621-976-2813 Sequence 2813, Ap
C 17	41.6	3.4	280	3	US-09-662-254B-27 Sequence 27, Appl
C 18	41.6	3.4	32392	3	US-09-662-254B-24 Sequence 24, Appl
C 19	41.2	3.4	50000	3	US-09-662-254B-24 Sequence 24, Appl
C 20	41	3.4	244	3	US-09-621-976-484 Sequence 484, Appl
C 21	41	3.4	19124	2	US-08-487-826B-13 Sequence 13, Appl
C 22	40.6	3.3	601	3	US-09-949-016-70227 Sequence 70227, A
C 23	40.6	3.3	70323	3	US-09-949-016-17594 Sequence 17594, A
C 24	40.4	3.3	837	3	US-08-998-416-288 Sequence 288, App

C 25	40.4	3.3	1412	3	US-09-270-767-7491 Sequence 7491, Ap
C 26	40.4	3.3	1412	3	US-09-270-767-22773 Sequence 22773, A
C 27	40.4	3.3	6996	3	US-09-573-080A-29 Sequence 29, Appl
C 28	40.4	3.3	278866	3	US-09-949-016-13922 Sequence 13922, A
C 29	40.4	3.3	278866	3	US-09-949-016-13923 Sequence 13923, A
C 30	40.4	3.3	278866	3	US-09-949-016-13924 Sequence 13924, A
C 31	40.4	3.3	278866	3	US-09-949-016-13925 Sequence 13925, A
C 32	40.4	3.3	278866	3	US-09-949-016-13926 Sequence 13926, A
C 33	40.4	3.3	278866	3	US-09-949-016-14699 Sequence 14699, A
C 34	40.4	3.3	278866	3	US-09-949-016-14700 Sequence 14700, A
C 35	40.4	3.3	278866	3	US-09-949-016-14701 Sequence 14701, A
C 36	40.4	3.3	278866	3	US-09-949-016-14702 Sequence 14702, A
C 37	40.4	3.3	278866	3	US-09-949-016-14703 Sequence 14703, A
C 38	40.2	3.3	601	3	US-09-949-016-206713 Sequence 206713, A
C 39	40.2	3.3	11748	2	US-08-611-107-30 Sequence 30, Appl
C 40	40	3.3	663	3	US-08-998-416-191 Sequence 191, App
C 41	40	3.3	921	3	US-09-248-796A-9583 Sequence 9583, Ap
C 42	40	3.3	7286	3	US-09-331-581-13 Sequence 13, Appl
C 43	40	3.3	7938	3	US-09-331-581-14 Sequence 14, Appl
C 44	40	3.3	72278	3	US-09-949-016-16113 Sequence 16113, A
C 45	39.8	3.3	19124	2	US-08-487-826B-13 Sequence 13, Appl

ALIGNMENTS

RESULT 1\nUS-08-232-463-14/c\n; Sequence 14, Application US/08232463\n; Patent No. 5670367\n\nGENERAL INFORMATION:\n; APPLICANT: DORNER, F.\n; APPLICANT: SCHIEFLINGER, F.\n; APPLICANT: FALKNER, F. G.\n; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS\n; NUMBER OF SEQUENCES: 52\n; CORRESPONDENCE ADDRESS:\n; ADDRESSEE: Foley & Lardner\n; STREET: 1800 Diagonal Road, Suite 500\n; CITY: Alexandria\n; STATE: VA\n; COUNTRY: USA\n; ZIP: 22313-0299\n\nCOMPUTER READABLE FORM:\n; MEDIUM TYPE: Floppy disk\n; COMPUTER: IBM PC compatible\n; OPERATING SYSTEM: PC-DOS/MS-DOS\n; SOFTWARE: Patent Release #1.0, Version #1.25\n\nCURRENT APPLICATION DATA:\n; APPLICATION NUMBER: US/08/232,463\n; FILING DATE:\n; CLASSIFICATION: 435\n; PRIOR APPLICATION DATA:\n; APPLICATION NUMBER: US/07/935,313\n; FILING DATE:\n; APPLICATION NUMBER: EP 91 114 300.6\n; FILING DATE: 26-AUG-1991\n; ATTORNEY/AGENT INFORMATION:\n; NAME: BENT, Stephen A.\n; REGISTRATION NUMBER: 29,768\n; REFERENCE/DOCKET NUMBER: 30472/114 IMMU\n; TELECOMMUNICATION INFORMATION:\n; TELEPHONE: (703) 836-9300\n; TELEFAX: (703) 683-4109\n; TELEX: 899149\n\nINFORMATION FOR SEQ ID NO: 14:\n; SEQUENCE CHARACTERISTICS:\n; LENGTH: 7218 base pairs\n; TYPE: nucleic acid\n; STRANDEDNESS: single\n; TOPOLOGY: linear\n; IMMEDIATE SOURCE:\n; CLONE: PTzapc-P15

NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele

LOCATION: 12429
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492
OTHER INFORMATION: 10-507-321 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 13524
OTHER INFORMATION: 10-507-353 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13535

Query Match
Best Local Similarity 49.0%; Pred. No. 0.038;
Matches 180; Conservative 0; Mismatches 181; Indels 6; Gaps 2;

QY 276 ATACTTCAGTCGAAATATCTATTTGTTACTTACAGACATTTATCAGGCCAAACCTTTA 335
DB 11439 AAAATTTAATTTAATTTAATTTGAAGCAATTTAAATTTAAATTTAATTTAATTTA 11380

QY 336 AGTTTATTTGATAGTATTTTGTGTGTTAAGCACAGACATGTGAGTCTGAGAAACTGA 395
DB 11379 AATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11320

QY 396 TGTGTGTAAGTCTGATTTAATA--ATATCAAACTGGTTAAATTAATTAATTAATTAATTAATTA 453
DB 11319 TTAATTTAAGTTAATTAATTAATTTAAATTTAATTTAATTTAATTTAATTTAATTTA 11260

QY 454 ACTTAAAAAAACCAATTCACAAAGACTT---ACTGCCCTGGAAAAAGCACAT 509
DB 11259 AATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTAATTTAATTTA 11200

QY 510 TTACAAGGCTGATCTAGCTTATCTATTTATTAACATTAAGAAATTAACAGAGTCCGAT 569
DB 11199 TAAATTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 11140

QY 570 ATAGCTGAATTTATTTCTTAAAGAAACAGAACTTCACTATCTTTAAATTAAGTG 629
DB 11139 AAAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 11080

QY 630 ATTACCT 636
DB 11079 CTTAGCT 11073

RESULT 5
US-10-170-097-651/C
Sequence 651, Application US/10170097
Patent No. 6794143
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
APPLICANT: Cohen, Annick
TITLE OR INVENTION: BIALLERIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
FILE REFERENCE: GEN-T114XC2D1
CURRENT APPLICATION NUMBER: US/10/170,097
PRIOR FILING DATE: 2002-06-10
PRIOR APPLICATION NUMBER: US 09/641,638
PRIOR FILING DATE: 2000-08-16
PRIOR APPLICATION NUMBER: US 09/502,330
PRIOR FILING DATE: 2000-02-11
PRIOR APPLICATION NUMBER: US 60/133,200
PRIOR FILING DATE: 1999-05-07
PRIOR APPLICATION NUMBER: US 09/275,267
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: US 60/119,917
PRIOR FILING DATE: 1999-02-12
NUMBER OF SEQ ID NOS: 1304
SOFTWARE: Patent.pm

```

; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; FEATURE:
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12854..13023
; OTHER INFORMATION: exon 10
; FEATURE:
; NAME/KEY: exon
; LOCATION: 13308..13429
; OTHER INFORMATION: exon 11
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16567..16667
; OTHER INFORMATION: exon 12
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16775..16945
; OTHER INFORMATION: exon 13
; FEATURE:
; NAME/KEY: exon
; LOCATION: 17063..17554
; OTHER INFORMATION: exon 14
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 17555..20674
; OTHER INFORMATION: 3'regulatory region
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1128
; OTHER INFORMATION: 10-508-191 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1182
; OTHER INFORMATION: 10-508-245 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1559
; OTHER INFORMATION: 10-509-284 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1570
; OTHER INFORMATION: 10-509-295 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 1827
; OTHER INFORMATION: 10-510-173 : variable motif ATTAA or TTTTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2048
; OTHER INFORMATION: 10-511-62 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2323
; OTHER INFORMATION: 10-511-337 : insertion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2341
; OTHER INFORMATION: 10-512-36 : polymorphic base G or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2623
; OTHER INFORMATION: 10-512-318 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2832
; OTHER INFORMATION: 10-513-250 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2844
; OTHER INFORMATION: 10-513-262 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2934
; OTHER INFORMATION: 10-513-352 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 2947
; OTHER INFORMATION: 10-513-365 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 3802
; OTHER INFORMATION: 12-206-81 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4062
; OTHER INFORMATION: 10-343-231 : deletion of C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4088
; OTHER INFORMATION: 12-206-366 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4109
; OTHER INFORMATION: 10-343-278 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 4170
; OTHER INFORMATION: 10-343-339 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 5903
; OTHER INFORMATION: 10-346-23 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
```

```
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 6429
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
```

```
Query Match 3.7%; Score 45.4; DB 3; Length 20674;
Best Local Similarity 49.0%; Pred. No. 0.038;
Matches 180; Conservative 0; Mismatches 181; Indels 6; Gaps 2;

Qy 276 ATACTCAGCTCAATATCTATTTGTTAACAAGACATTAATCAGCCAAACCACTTTA 335
Db 11439 AATATTTAATTTATATATGAGCAATTAATAATTAATTAATTAATTAATTA 11380

Qy 336 AGTTTATTTGTAATGATTTTGTGTTAAGCAGACATGAGTGTGAGAAACTGA 395
Db 11379 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11320

Qy 396 TGTGGTAACTGATTTAATA--ATATCAACTGGTTAAATTAATAAATTAATTA 453
Db 11319 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11260

Qy 454 ACTTAAAAAACAATTAACCAACACACTT---ACTTCCCTTGAAAAACACAT 509
Db 11259 ATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 11200

Qy 510 TTACAAGGCTGATCTTACCTTTAATTTAATTAAGAAATTAACCAAGTCCGAT 569
Db 11199 TAAATTTAATTAATTTAATTAATTAATTAATTAATTAATTAATTAATTA 11140

Qy 570 ATAGCTGTAATTTATCTTAAGACAGAACTTCACTATGCTTTAAATTAAGTG 629
Db 11139 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC 11080

Qy 630 ATTACCT 636
Db 11079 CTTAGCT 11073
```

```
RESULT 6
US-08-998-416-1036/c
; Sequence 1036, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Rohmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
```

```
APPLICANT: Wendland, Jürgen
APPLICANT: Knechtle, Philipp
APPLICANT: Redischung, Corinne
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPYII
TITLE OF INVENTION: AND USES THEREOF
NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSER: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1036:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAG1633UP
US-08-998-416-1036
```

```
Query Match 3.7%; Score 44.6; DB 3; Length 732;
Best Local Similarity 46.3%; Pred. No. 0.01;
Matches 146; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

Qy 339 TTTATTTGTAATTTTGTGTTAAGCAGACATGTGAGTGTGAGAAACGATGT 398
Db 732 TCTATATTAATTAATGTTTATCTTAATTAATTAATTAATTAATTAATTA 673

Qy 399 TGTGAATCTGATTTAATATCAAACTGGTTAAATTAATAAATGATGATTA 458
Db 672 TTTAATTAATGATTAATTAATCAAAAAATTAATTAATTAATTAATTAATTA 613

Qy 459 AAAAAAACAATTAACAAGACTTACTTCCCTTGAAAAAGACATTTACAAGG 518
Db 612 CTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 553

Qy 519 CTGATCTTACCTTTATTTAATTAACAATTAAGAAATTAACCAAGTCCGAT 578
Db 552 AATTAATTTATGATTTAATTAATTAATTAATTAATTAATTAATTAATTA 493

Qy 579 ATTTATTCCTTAAGACAGAACTTCACTATGCTTTAAATTAAGTGAATCA 638
Db 492 AGTAATATCAATTTAAAGGTAGACTATTAATTAAGAAATTAATTAATTA 433

Qy 639 GATACCTGCAAGCT 653
Db 432 AATATTAATTAAGGT 418
```

RESULT 7

Query Match	3.6%;	Score 44;	DB 3;	Length 1141;
Best Local Similarity	10.7%;	Pred. No. 0.02;		
Matches 54;	Conservative 227;	Mismatches 215;	Indels 11;	Gaps 2

RESULT 8
US-10-104-047-682/c

```

:
: CURRENT FILING DATE: 2002-03-22
:
: PRIOR APPLICATION NUMBER:
:
: PRIOR FILING DATE:
:
: NUMBER OF SEQ ID NOS: 4096
:
: SOFTWARE: PatentIn Ver. 2.1
:
: SEQ ID NO 682
:
: LENGTH: 2308
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: US-10-104-047-682

```

Query Match	3.6%	Score 43.8	DB 3	Length 2308
Best Local Similarity	47.6%	Pred. No. 0.033		
Matches 129; Conservative	0	Mismatches 142	Indels 0	Gaps 0

Qy 653 TTAGCTACAAACGAGCAGACAGACAACAGA 683
Db 1352 AAACCTTGAGGTAGACTAAGAAAAAACAAGA 1322

! OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15777

Query Match	3.5%;	Score 42.6;	DB 3;	Length 80411;
Best Local Similarity	52.5%;	Pred. No. 0.55;	Matches 99;	Conservative 0;
Mismatches	0;	Mismatches	84;	Indels 0;
Gaps	0;	Gaps	0;	
Qy	305	AACAGACATTAATTCAGGCCAAACCCCTTAAGTTTATTTGATAGTATTTTGCTGTAA	364	
Db	17141	AAATGACATCTTTTAATTCCTTATCCAGAGTAAAGTACATTTATGATATCGATGATAGCTTAT	17083	

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 12, 2005, 12:05:36 ; Search time 996 Seconds

(without alignments)
10129.158 Million cell updates/sec

Title: US-10-627-075-1

Perfect score: 1220

Sequence: 1 gagctctccagctaccgacgc.....ccaaaagagagagagatg 1220

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA_Main:*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220	100.0	1220	7	US-10-627-075-1
2	276	22.6	2475	9	US-10-764-420-1571
3	238	19.5	240825	3	US-09-790-289-1
4	238	19.5	240825	7	US-10-468-582-1
5	238	19.5	240825	10	US-11-028-971-1
6	213	17.5	2424	7	US-10-317-803-11
7	213	17.5	2424	8	US-10-789-222-34
8	213	17.5	2424	8	US-10-827-759A-5
9	211.8	17.4	62705	7	US-10-317-803-4
10	156.6	12.8	3251	5	US-10-044-090-24
11	140.6	11.5	2846	7	US-10-433-793-187
12	139.4	11.4	2846	7	US-10-433-793-188
13	55	4.5	15698	6	US-10-311-455-866
14	54.4	4.5	15698	6	US-10-311-455-2114
15	52.6	4.3	2265	8	US-10-602-494-272
16	51.6	4.2	6823	6	US-10-311-455-1117
17	51.2	4.2	20486	6	US-10-240-485-164
18	51	4.2	3673778	6	US-10-312-841-2
19	50.2	4.1	9964	6	US-10-311-455-72
20	50	4.1	4415	6	US-10-311-455-1879
21	50	4.1	4415	6	US-10-240-485-155
22	49.8	4.1	15479	7	US-10-257-166-45
23	49.2	4.0	13784	7	US-10-257-166-144

C	24	49	4.0	6127	6	US-10-240-485-2	Sequence 2, Appl
C	25	49	4.0	8758	6	US-10-311-455-1091	Sequence 1091, Ap
C	26	48.6	4.0	487	3	US-09-918-995-5922	Sequence 5922, Ap
C	27	48.4	4.0	4661	7	US-10-221-613-93	Sequence 93, Appl
C	28	48.2	4.0	2265	8	US-10-602-494-130	Sequence 130, Appl
C	29	48	3.9	502	3	US-09-814-353-17606	Sequence 17606, A
C	30	48	3.9	5919	6	US-10-311-455-362	Sequence 362, App
C	31	48	3.9	5919	6	US-10-221-613-64	Sequence 64, Appl
C	32	48	3.9	8056	6	US-10-311-455-1738	Sequence 1738, Ap
C	33	48	3.9	8056	6	US-10-473-126-386	Sequence 386, App
C	34	48	3.9	8895	6	US-10-311-455-305	Sequence 305, App
C	35	47.8	3.9	14950	6	US-10-311-455-1229	Sequence 1229, Ap
C	36	47.8	3.9	16724	6	US-10-311-455-1063	Sequence 1063, Ap
C	37	47.8	3.9	16724	6	US-10-240-485-89	Sequence 89, Appl
C	38	47.6	3.9	1478	5	US-10-198-846-10387	Sequence 10387, A
C	39	47.6	3.9	6103	6	US-10-311-455-1664	Sequence 1664, Ap
C	40	47.6	3.9	8056	8	US-10-473-126-240	Sequence 240, App
C	41	47.6	3.9	3673778	6	US-10-312-841-1	Sequence 1, Appl
C	42	47.4	3.9	302	3	US-09-814-353-5335	Sequence 5335, Ap
C	43	47.4	3.9	302	3	US-09-814-353-11622	Sequence 11622, A
C	44	47.2	3.9	545	4	US-09-925-065A-313319	Sequence 313319,
C	45	47.2	3.9	2240	8	US-10-473-126-102	Sequence 102, App

ALIGNMENTS

RESULT 1	US-10-627-075-1
1	Sequence 1, Application US/10627075
2	Publication No. US20040091913A1
3	GENERAL INFORMATION:
4	APPLICANT: Livingston et al.
5	TITLE OF INVENTION: Composition and Method for Imaging Cells
6	FILE REFERENCE: 20363-019
7	CURRENT APPLICATION NUMBER: US/10/627,075
8	PRIOR FILING DATE: 2003-07-24
9	PRIOR APPLICATION NUMBER: 60/398,583
10	PRIOR FILING DATE: 2002-07-25
11	NUMBER OF SEQ ID NOS: 1
12	SOFTWARE: PatentIn Ver. 2.1
13	SEQ ID NO 1
14	LENGTH: 1220
15	TYPE: DNA
16	ORGANISM: Mus musculus
17	US-10-627-075-1
Query Match	100.0%; Score 1220; DB 7; Length 1220;
Best Local Similarity	100.0%; Pred. No. 2.4e-308; Indels 0; Gaps 0;
Matches 1220; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 GAGCTTCCAGTACCGATCTGCGAGCTTAACTTCTAGTCATGAAGGGTGTGACTC 60
DB	1 GAGCTTCCAGTACCGATCTGCGAGCTTAACTTCTAGTCATGAAGGGTGTGACTC 60
QY	61 TGGACGAGAGAGCCAGAGCTGAGAGTGTGTTAGAGTCAGTCAGGCCAGCTTTT 120
DB	61 TGGACGAGAGAGCCAGAGCTGAGAGTGTGTTAGAGTCAGTCAGGCCAGCTTTT 120
QY	121 ATGGCCAGAGGGCTTTGAACTTAATTAAGGGAAGTATTTGCTAGCCCACTGAC 180
DB	121 ATGGCCAGAGGGCTTTGAACTTAATTAAGGGAAGTATTTGCTAGCCCACTGAC 180
QY	181 TGGAGTAATTAATCAGGAATGCGACAGAGTATGAGCCCGAGGAACCTGTATCA 240
DB	181 TGGAGTAATTAATCAGGAATGCGACAGAGTATGAGCCCGAGGAACCTGTATCA 240
QY	241 GTGAAGAAAGGTGTATGTTTCTCTGACATACCTCAATATCTATGTT 300
DB	241 GTGAAGAAAGGTGTATGTTTCTCTGACATACCTCAATATCTATGTT 300
QY	301 ACTTAACACACATTAATGAGCCAAACCTTTAAGTTTATTTGATATTTTGTG 360
DB	301 ACTTAACACACATTAATGAGCCAAACCTTTAAGTTTATTTGATATTTTGTG 360

Db 301 ACTTAAGACAACTTAATCAAGCCAAACCACTTTAAGTTTATTTGATATTTGTC 360
Qy 361 TTAAGGACACAGACTGTGAGTCTGAGAAACTGATGTTGGTAACTTGATTTATATAT 420
Db 361 TTAAGGACACAGACTGTGAGTCTGAGAAACTGATGTTGGTAACTTGATTTATATAT 420
Qy 421 CAAACGGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
Db 421 CAAACGGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
Qy 481 AGACTTACTTCCCTTGAAGAACCACTTACAAAGGCTGATCTTACCTTTATATTTA 540
Db 481 AGACTTACTTCCCTTGAAGAACCACTTACAAAGGCTGATCTTACCTTTATATTTA 540
Qy 541 CAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Db 541 CAATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
Qy 601 AACTTCACTATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
Db 601 AACTTCACTATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 660
Qy 661 CAAACGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 720
Db 661 CAAACGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 720
Qy 721 CTCTGACATGCGGAGGCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Db 721 CTCTGACATGCGGAGGCTTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy 781 CAGTCTTTGGGCGAGTAACTATGCTGTGATTTTCTGTGTGTGTGTGTGTGTGTGTGTGT 840
Db 781 CAGTCTTTGGGCGAGTAACTATGCTGTGATTTTCTGTGTGTGTGTGTGTGTGTGTGTGT 840
Qy 841 CCTACGAGAAATAGT 900
Db 841 CCTACGAGAAATAGT 900
Qy 901 CTCTTATCACTATATATTAATTAAGGAAAGAAATGATTTGATTTGATTAATTAATTAAT 960
Db 901 CTCTTATCACTATATATTAATTAAGGAAAGAAATGATTTGATTTGATTAATTAATTAAT 960
Qy 961 ACTCAGGAGGAGAAACAAAGAGTCCGTGACAGACCTCTGAGTGAAGCAGGCTCTCTTCC 1020
Db 961 ACTCAGGAGGAGAAACAAAGAGTCCGTGACAGACCTCTGAGTGAAGCAGGCTCTCTTCC 1020
Qy 1021 TCTCAGGACAGCTCCGAGTGTGCTCGGAGGAGAAAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TCTCAGGACAGCTCCGAGTGTGCTCGGAGGAGAAAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 GAGCCTGCTGCGGAGGAG 1140
Db 1081 GAGCCTGCTGCGGAGGAG 1140
Qy 1141 GTGCTTTAGACAGCAGCTGAGAGCTCAGAGCCAAAGTTTGTGAATCACTACATTTAGAAC 1200
Db 1141 GTGCTTTAGACAGCAGCTGAGAGCTCAGAGCCAAAGTTTGTGAATCACTACATTTAGAAC 1200
Qy 1201 CCAAAAAG 1220
Db 1201 CCAAAAAG 1220

RESULT 2
US-10-764-420-1571
; Sequence 1571, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yefun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent Possesses A Defined Biological Activity

FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1571
; LENGTH: 2475
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1571

Query Match 22.6%; Score 276; DB 9; Length 2475;
Best Local Similarity 100.0%; Pred. No. 4,3e-61;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 945 GATACTGACACTGTAGACTCAGGAGGAGAAACAAAGAGTCCGTGACAGACCTCTGAGAGT 1004
Db 1 GATACTGACACTGTAGACTCAGGAGGAGAAACAAAGAGTCCGTGACAGACCTCTGAGAGT 60
Qy 1005 CAGGCTGCTCTTCTCTCTCAGACAGCTCCGAGTGTCCCGGAGAGAGAGAGAGAGAG 1064
Db 61 CAGGCTGCTCTTCTCTCTCAGACAGCTCCGAGTGTCCCGGAGAGAGAGAGAGAGAGAG 120
Qy 1065 GACAGGACACTGAGAGAGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1124
Db 121 GACAGGACACTGAGAGAGAGCTGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy 1125 CACAGGACAGAGCTGCTGCTTGAACAGCAGCTGAGAGCTCAGAGAGCAAGTTTGCTGA 1184
Db 181 CACAGGACAGAGCTGCTGCTTGAACAGCAGCTGAGAGCTCAGAGAGCAAGTTTGCTGA 240
Qy 1185 ACTCAGCTTTGAAG 1220
Db 241 ACTCAGCTTTGAAG 276

RESULT 3
US-09-790-289-1/c
; Sequence 1, Application US/09790289
; Publication No. US20030165826A1
; GENERAL INFORMATION:
; APPLICANT: Caroline Barry
; APPLICANT: Ilya Chumakov
; TITLE OF INVENTION: PG-3 and Biallelic Markers Thereof
; FILE REFERENCE: 68, US3, REG
; CURRENT APPLICATION NUMBER: US/09/790,289
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 240825
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2000
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 2001..2079
; OTHER INFORMATION: exon A
; NAME/KEY: exon
; LOCATION: 4627..4718
; OTHER INFORMATION: exon B
; NAME/KEY: exon
; LOCATION: 10115..10233
; OTHER INFORMATION: exon C
; NAME/KEY: exon
; LOCATION: 26810..26897
; OTHER INFORMATION: exon D

NAME/KEY: exon
LOCATION: 31357..31471
OTHER INFORMATION: exon E
NAME/KEY: exon
LOCATION: 34261..34404
OTHER INFORMATION: exon F
NAME/KEY: exon
LOCATION: 37377..37466
OTHER INFORMATION: exon S
NAME/KEY: exon
LOCATION: 39704..40858
OTHER INFORMATION: exon T
NAME/KEY: exon
LOCATION: 50436..50545
OTHER INFORMATION: exon G
NAME/KEY: exon
LOCATION: 72881..72918
OTHER INFORMATION: exon H
NAME/KEY: exon
LOCATION: 75989..76151
OTHER INFORMATION: exon I
NAME/KEY: exon
LOCATION: 95111..95188
OTHER INFORMATION: exon J
NAME/KEY: exon
LOCATION: 216015..216252
OTHER INFORMATION: exon K
NAME/KEY: exon
LOCATION: 237526..238825
OTHER INFORMATION: exon L
NAME/KEY: misc feature
LOCATION: 238826..240825
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1999
OTHER INFORMATION: 5-390-177 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 4601
OTHER INFORMATION: 5-391-43 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 10228
OTHER INFORMATION: 5-392-222 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 10286
OTHER INFORMATION: 5-392-280 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 10370
OTHER INFORMATION: 5-392-364 : insertion of G
NAME/KEY: allele
LOCATION: 39944
OTHER INFORMATION: 4-58-318 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 39973
OTHER INFORMATION: 4-58-289 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 41385
OTHER INFORMATION: 4-54-199 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 41404
OTHER INFORMATION: 4-54-180 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 42332
OTHER INFORMATION: 4-51-312 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 67475
OTHER INFORMATION: 99-86-266 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 69521
OTHER INFORMATION: 4-88-107 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 72838
OTHER INFORMATION: 5-397-141 : polymorphic base G or T
NAME/KEY: allele

LOCATION: 76060
OTHER INFORMATION: 5-398-203 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 81253
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 83921
OTHER INFORMATION: 99-109-358 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 91917
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95349
OTHER INFORMATION: 4-21-154 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 95511
OTHER INFORMATION: 4-21-317 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 96190
OTHER INFORMATION: 4-23-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 97294
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 98024
OTHER INFORMATION: 5-364-252 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 98914
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 98963
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 103593
OTHER INFORMATION: 4-87-212 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 104398
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 106373
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 106407
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 108315
OTHER INFORMATION: 4-105-98 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 108327
OTHER INFORMATION: 4-105-86 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 108472
OTHER INFORMATION: 4-45-49 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 109196
OTHER INFORMATION: 4-44-277 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 114604
OTHER INFORMATION: 4-86-60 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 115716
OTHER INFORMATION: 4-84-334 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 122083
OTHER INFORMATION: 99-78-321 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 123124
OTHER INFORMATION: 99-12767-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 123231
OTHER INFORMATION: 99-12767-143 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 123277

```
OTHER INFORMATION: 99-12767-189 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 123468
OTHER INFORMATION: 99-12767-380 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 126738
OTHER INFORMATION: 4-80-328 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 128210
OTHER INFORMATION: 4-36-384 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 128330
OTHER INFORMATION: 4-36-264 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 128333
OTHER INFORMATION: 4-36-261 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 128594
OTHER INFORMATION: 4-35-333 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 128687
OTHER INFORMATION: 4-35-240 : polymorphic base G or C
NAME/KEY: allele
```

```
Query Match 19.5%; Score 238; DB 3; Length 240825;
Best Local Similarity 64.3%; Pred. No. 4,7e-50;
Matches 545; Conservative 0; Mismatches 265; Indels 38; Gaps 11;
```

```
QY 221 CCCGAGAAACCTGATACAGTAAAGAAAGGTGATGTTGTTCTCTCGACATACCT 280
DB 158572 CTCGAAAAAAGGTTGATTCCTCTTCTACAGACACTT 158513
QY 281 TCACGCAATATCTATGTTACTTACAGCATTAATCGGCAACCACTTAAGTTT 340
DB 158512 AAATTTAATTTCTTATGTAATGTAAGTAACTTAATGACAGTATTTTGAAGT 158453
QY 341 TATTTGATAGTATTTTGTGAAGCACAGCATGTGAGTGTGAGAAATGATGTTG 400
DB 158452 GAATATCCAAAGTGAAGTTTTTAAAGA-ATAAAGTCAGAGCTCAG-AACTGATGTT 158395
QY 401 GTAACCTGATTTAATA-----TATCAACTGGGTTAAATAAAA-----AA 443
DB 158394 CTAACTGATTTAATAAGATCTTTCTATCCAACTGTCGAAATGAATACTTTTGT 158335
QY 444 AATGTCATTAACCTTAATAAAAAAACAATAACCAAGACTTTACTCCCTGGAATA 503
DB 158334 ACTTGCTAGTGTGTCAGAAATCACTAAATCAAGATCTATTTTCTCTGAAAG 158275
QY 504 GCACATTTCAAGGG--CTGATCTTAGCCTTTATTTTACATAAAGAAATAAACCA 560
DB 158274 GCACAGTTTACAGAGGCCAACTCTCTTTATTTGTAATAAATAAACCA 158215
QY 561 GGTCCGATATAGCTGTATTTTATTCCTA-----AAAGAACGAAACTTTCATG 613
DB 158214 AGTCTGACCTATTTGTATTTTATTTTCTTAAAGAAAAACGAACTTTCATGTA 158155
QY 614 CTTTAAATTAAGTATTAACCTCAAGTACTGCAAGCTTGAAGCTTACAAAGACAG-- 671
DB 158154 CTTCAACATTAAGTTATTAACCTCAGATATTTTGCAGCTTACAGGCAAAATCACT 158095
QY 672 ACAGCAACAGAGCCCA--GCTACTCTCTAGGAAA---TAATTAGGGTGTGCTCTGA 726
DB 158094 TCAGACAAAGAGATCACTGCTCTCTTAGGAAATACTTAATGGGGTGTGCTTAGA 158035
QY 727 CATGCCAGAGGCTTGTGCTGTGCTG--TGTTCCAGAGAGCTTCTCAGATCACTC 785
DB 158034 AATGCCAGAGGCTCTGTAAAGATCGGTTTTTCCAGAGGCTTCTCAGATGAGGCTC 157975
QY 786 CTTTGGGAGTAACTAGCTGATTTTCTGTTGCTGCTGCTAGTACCCCTAC 845
DB 157974 TGGTTGAGGAGGAGCATTTGCTGATTTTCTGTTGCTGCTGCTAGTACCCCTAC 157915
QY 846 AGAAGATATGTGGTGAAGCAGG--GGGGGAGCGGCTGCTGACATGTCTGGCTGCT 904
```

```
DB 157914 AGAAGATATCGGCTTAAGCAGAGGCGGAGACCCCACTACATGTTCTGCTCT 157855
QY 905 TATCAACTATCATATTAAGGAAAGAAAGTATGATTTGGATTAAGTACTGACCTGTA 964
DB 157854 TATCAACTATCATATTAAGGAAAGAAAGTATGATTTGGATTAAGTACTGACCTGTA 157795
QY 965 AGGAGAAACAAAGAGCTCGTGAACCTCTGAGTGAAGCAGGCTGCTCTTCTCTC 1024
DB 157794 TGGGAGAGAGAGAAACAAAGACCTGTGAAGCTGCTCTGTAAAGCTGACACAGCCCTCC 157735
QY 1025 AGGACAGC 1032
DB 157734 AAGTGAGC 157727

RESULT 4
US-10-468-582-1/c
Sequence 1, Application US/10468582
Publication No. US20040163137A1
GENERAL INFORMATION:
APPLICANT: Genet
TITLE OF INVENTION: PG-3 and diallelic markers thereof
FILE REFERENCE: 68,WO2
CURRENT APPLICATION NUMBER: US/10/468,582
CURRENT FILING DATE: 2003-08-20
NUMBER OF SEQ ID NOS: 5
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 240825
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..2000
OTHER INFORMATION: 5' regulatory region
FEATURE:
NAME/KEY: exon
LOCATION: 2001..2079
OTHER INFORMATION: exon A
FEATURE:
NAME/KEY: exon
LOCATION: 4627..4718
OTHER INFORMATION: exon B
FEATURE:
NAME/KEY: exon
LOCATION: 10115..10233
OTHER INFORMATION: exon C
FEATURE:
NAME/KEY: exon
LOCATION: 26810..26897
OTHER INFORMATION: exon D
FEATURE:
NAME/KEY: exon
LOCATION: 31357..31471
OTHER INFORMATION: exon E
FEATURE:
NAME/KEY: exon
LOCATION: 34261..34404
OTHER INFORMATION: exon F
FEATURE:
NAME/KEY: exon
LOCATION: 37377..37466
OTHER INFORMATION: exon S
FEATURE:
NAME/KEY: exon
LOCATION: 39704..40858
OTHER INFORMATION: exon T
FEATURE:
NAME/KEY: exon
LOCATION: 50436..50545
OTHER INFORMATION: exon G
FEATURE:
```


NAME/KEY: exon
LOCATION: 72881..72918
OTHER INFORMATION: exon H
FEATURE:
NAME/KEY: exon
LOCATION: 75989..76151
OTHER INFORMATION: exon I
FEATURE:
NAME/KEY: exon
LOCATION: 95111..95188
OTHER INFORMATION: exon J
FEATURE:
NAME/KEY: exon
LOCATION: 216015..216252
OTHER INFORMATION: exon K
FEATURE:
NAME/KEY: exon
LOCATION: 237526..238825
OTHER INFORMATION: exon L
FEATURE:
NAME/KEY: misc_feature
LOCATION: 238826..240825
OTHER INFORMATION: 3'regulatory region
FEATURE:
NAME/KEY: allele
LOCATION: 1999
OTHER INFORMATION: 5-390-177 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 4601
OTHER INFORMATION: 5-391-43 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 10228
OTHER INFORMATION: 5-392-222 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 10286
OTHER INFORMATION: 5-392-280 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 10370
OTHER INFORMATION: 5-392-364 : insertion of G
FEATURE:
NAME/KEY: allele
LOCATION: 39973
OTHER INFORMATION: 4-58-289 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 41385
OTHER INFORMATION: 4-54-199 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 41404
OTHER INFORMATION: 4-54-180 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 42232
OTHER INFORMATION: 4-51-312 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 67475
OTHER INFORMATION: 99-86-266 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 69521
OTHER INFORMATION: 4-88-107 : polymorphic base A or G
FEATURE:
NAME/KEY: allele

LOCATION: 72838
OTHER INFORMATION: 5-397-141 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 76060
OTHER INFORMATION: 5-398-203 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 81253
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 83921
OTHER INFORMATION: 99-109-358 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 91917
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95349
OTHER INFORMATION: 4-21-154 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95511
OTHER INFORMATION: 4-21-317 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 96190
OTHER INFORMATION: 4-23-326 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 97294
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 98024
OTHER INFORMATION: 5-364-252 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 98914
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 98963
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 103593
OTHER INFORMATION: 4-87-212 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 104398
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 106373
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 106407
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 108315
OTHER INFORMATION: 4-105-98 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108327
OTHER INFORMATION: 4-105-86 : polymorphic base A or G
FEATURE:
NAME/KEY: allele

LOCATION: 23825..24025
OTHER INFORMATION: 3'regulatory region
FEATURE:
NAME/KEY: allele
LOCATION: 1999
OTHER INFORMATION: 5-390-177 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 4601
OTHER INFORMATION: 5-391-43 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 10228
OTHER INFORMATION: 5-392-222 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 10286
OTHER INFORMATION: 5-392-280 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 10370
OTHER INFORMATION: 5-392-364 : insertion of G
FEATURE:
NAME/KEY: allele
LOCATION: 39944
OTHER INFORMATION: 4-58-318 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 39973
OTHER INFORMATION: 4-58-289 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 41385
OTHER INFORMATION: 4-54-199 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 41404
OTHER INFORMATION: 4-54-180 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 42232
OTHER INFORMATION: 4-51-312 : polymorphic base G or C
FEATURE:
NAME/KEY: allele
LOCATION: 67475
OTHER INFORMATION: 99-86-266 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 69521
OTHER INFORMATION: 4-88-107 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 72838
OTHER INFORMATION: 5-397-141 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 76060
OTHER INFORMATION: 5-398-203 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 81253
OTHER INFORMATION: 99-12738-248 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 83921
OTHER INFORMATION: 99-109-358 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 91917
OTHER INFORMATION: 99-12749-175 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95349

OTHER INFORMATION: 4-21-154 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 95511
OTHER INFORMATION: 4-21-317 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 96190
OTHER INFORMATION: 4-23-326 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 97294
OTHER INFORMATION: 99-12753-34 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 98024
OTHER INFORMATION: 5-364-252 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 98914
OTHER INFORMATION: 99-12755-280 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 98963
OTHER INFORMATION: 99-12755-329 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 103593
OTHER INFORMATION: 4-87-212 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 104398
OTHER INFORMATION: 99-12757-318 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 106373
OTHER INFORMATION: 99-12758-102 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 106407
OTHER INFORMATION: 99-12758-136 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 108315
OTHER INFORMATION: 4-105-98 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108327

Query Match 19.5%; Score 238; DB 10; Length 240825;
Best Local Similarity 64.3%; Pred. No. 4,7e-50;
Matches 545; Conservative 0; Mismatches 265; Indels 38; Gaps 11;

QY 221 CCCGAGAAACCTGATACAGTGAAGAAAGCTGATGTTGTTCCCTCTGACATACT 280
DB 158572 CTCAAAAAAGGAAAAAAGGTTGATTCCTTTCTACAGACACTT 158513
QY 281 TCACCAATATCTATGTTTACCTTAACAGACATTATCAGGCCAAACCACTTAAGTT 340
DB 158512 AATTTCTAATTTCTCTATGTTAATGAACTTATATAGACATGATTTTGAAGT 158453
QY 341 TATTTGTATAGTATTTTGTGTTAAGCACAGACATGTGAGTCTGAGAAACTGATGTTG 400
DB 158452 GAATATCCAAAGTGAGTTTAAAAAGA-ATTAAGGTGAGACTCG-AAACTGATGTTT 158395
QY 401 GTAACCTGATTAATAA-----TATCAACTGGGTTAAATTAATAA-----AA 443
DB 158394 CTAACTGACTTAATAACAGATCTTCTATCCACTGTTGGAAGAAATCTTTTGTGA 158335
QY 444 AATGTCATTAACCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 503
DB 158334 ACTTGCTAGTGCTTCAAGAAATCAACTAATAATGATCTATTTTCCCTGGAAGAAG 158275


```
; PRIOR APPLICATION NUMBER: US 60/463,981
; PRIOR FILING DATE: 2003-04-18
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2308
; OTHER INFORMATION: n = a, g, c or t
US-10-827-759A-5

Query Match      17.5%; Score 213; DB 8; Length 2424;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1008 GGCTGCTCTCTCTCTCAGACAGCTCCAGTGTGCGCGGAGAAAGAAAGAGAGAC 1067
DB 1 GGCTGCTCTCTCTCTCAGACAGCTCCAGTGTGCGCGGAGAAAGAAAGAGAGAC 60
QY 1068 AGGCACTGGGAAAGAGCTCTGCGGAGCGAGAGAGCTCTCACTGATGACTTATTTCAC 1127
DB 61 AGGCACTGGGAAAGAGCTCTGCGGAGCGAGAGAGCTCTCACTGATGACTTATTTCAC 120
QY 1128 ACGGCAAGCCCTGTGCTTGAAGACAGAGCTGAGAGCTCAGAGCAAGTTTGTCTGAAT 1187
DB 121 ACGGCAAGCCCTGTGCTTGAAGACAGAGCTGAGAGCTCAGAGCAAGTTTGTCTGAAT 180
QY 1188 CACAGTTTGAACCCCAAAAAGAGAGAGAGATG 1220
DB 181 CACAGTTTGAACCCCAAAAAGAGAGAGAGATG 213

RESULT 9
US-10-317-803-4
; Sequence 4, Application US/10317803
; Publication No. US20040115640A1
; GENERAL INFORMATION:
; APPLICANT: Kathleen Myers
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: MODULATION OF ANGIOPOIETIN-2 EXPRESSION
; FILE REFERENCE: RTS-0454
; CURRENT APPLICATION NUMBER: US/10/317,803
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 244
; SEQ ID NO 4
; LENGTH: 62705
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20020137081A1 343965.2
US-10-317-803-4

Query Match      17.4%; Score 211.8; DB 7; Length 62705;
Best Local Similarity 74.3%; Pred. No. 1.6e-43;
Matches 335; Conservative 0; Mismatches 107; Indels 9; Gaps 5;

QY 591 AAAGAACAGAACTTCACTATGCTTAAATTAAGTATTAACCTCAGATCTAGCA 650
DB 31 AAAGAACAGAACTTCACTATGCTTAAATTAAGTATTAACCTCAGATCTAGCA 90
QY 651 GCTTACCTTCAAAACGAGAG--ACAGACAAACAGAGCCCCA--GCTACTCTTGAAGAA-- 705
DB 91 GCTTACCTTCAAAACGAGAG--ACAGACAAACAGAGCCCCA--GCTACTCTTGAAGAA-- 150
QY 706 --TAAATTAAGTGTGCTCTGCAATGCGGAGGCTTGTGCTGTGCTG--TGTTCACA 762
DB 151 CTTAAATTAAGTGTGCTCTGCAATGCGGAGGCTTGTGCTGTGCTG--TGTTCACA 210
QY 763 GAAGGCTTCTGAGTACAGAGTCTTGTGGGAGTAGACATATGCTGATTTTCTCG 822
DB 211 GAAGGCTTCTGAGTACAGAGTCTTGTGGGAGTAGACATATGCTGATTTTCTCG 270
```

```
QY 823 TTGCTGCTAGTAGACCCCTCAAGAGAAATAGTGGTAGACCGAG--GGCGGAGCGCT 881
DB 271 TTGCTGCTAGTAGACCCCTCAAGAGAAATAGTGGTAGACCGAG--GGCGGAGCGAC 330
QY 882 GGCTGACATGTGTGCTGCTCTTATCACTTATCATATTAAGGAAAGAAAGATTGAT 941
DB 331 CACTACACATGTGTGCTGCTCTTATCACTTATCATATTAAGGAAAGAAAGATTGAT 390
QY 942 TCGATACCTGACACTGTAGACTCAGGGAGAAACAAAGAGTCCGTCAGACCTCTGAGT 1001
DB 391 TCGATACCTGACACTGTAGACTCAGGGAGAAACAAAGAGTCCGTCAGACCTCTCT 450
QY 1002 GAGCAGGCTGCTCTCTCTCTCAGAGAC 1032
DB 451 GTAAAGCTGACACAGCCCTCCCAAGTAGAC 481

RESULT 10
US-10-044-090-24
; Sequence 24, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 3251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 343965.2
US-10-044-090-24

Query Match      12.8%; Score 156.6; DB 5; Length 3251;
Best Local Similarity 74.9%; Pred. No. 8.9e-30;
Matches 209; Conservative 0; Mismatches 69; Indels 1; Gaps 1;

QY 755 TGTTCACAGAGGCTTCTGACATACAGTCTTTGGGAGTAGACACTATGCTGTAT 814
DB 12 TTTCCAGAGGCTTCTGACATAGGCTGTGGTAGAGGAGGAGCAATTCTGTGAT 71
QY 815 TTTTCTGCTGCTGCTGCTGATGACCCCTTACAGAAAGATAGTGGTAGACCGAG--GGGG 873
DB 72 TTTTCTGCTGCTGCTGCTGATGACCCCTTACAGAAAGATAGTGGTAGACCGAGGGGG 131
QY 874 GAGCGGCTGCTGACATATGCTGCTGCTTATCACTTATCATATTAAGGAAAGAAAG 933
DB 132 GAGCGGCCACTTACATATGCTGCTGCTTATCACTTATCATATTAAGGAAAGAAAG 191
QY 934 TGAATTGATTGATACCTGACACTGTAGACTCAGGGAGAAACAAAGAGTCCGTCAGACC 993
DB 192 TGAATTGATTGATACCTGACACTGTAGACTCAGGGAGAGAGAAACAAAGAGTCCGTA 251
QY 994 TCTGAGTAGAGCAGGGCTGCTCTCTCTCTCAGAGAC 1032
DB 252 GCTGCTGTGTAAGCTGACAGACGCTCCCAAGTAGAC 290

RESULT 11
US-10-433-793-187
; Sequence 187, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
```

```

; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 187
; LENGTH: 2846
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-187

Query Match          11.5%; Score 140.6; DB 7; Length 2846;
Best Local Similarity 62.1%; Pred. No. 1.3e-25;
Matches 299; Conservative 0; Mismatches 164; Indels 16; Gaps 4;

QY 522 ATCTTACCTTATATTAATTAAGAAAATAAACAAGTCCGATATAGCTGTAT 581
DB 1530 ATTTTATTTTATTTATTTGTAATTAATAAATGAAGTTTGTATTTATTTGTA 1589
QY 582 TTAATTCCT-----AAAGAAGAAACTTTCACTATGCTTTAAATTAAGTGAT 634
DB 1590 TTTTATTTTAAAGAAAATAAGAAATTTATTTATTTAATTAATAAGTTATTA 1649
QY 635 CTCAGATCTCTCAAGCTTAGCCTTACAAACGAGACAGACAGACAGACCCAGCTAC 694
DB 1650 TTTAGATATTTTGTAGTTAGTACGGTAAATAATTGTTTATTAAGATTAATG 1709
QY 695 TCTCT-----AGGAAATAATTAGGTTGCTGCTCTGACATGCCAGGGGCTTGTGC 747
DB 1710 TTTTATTTAAGAAATTTTAAATTTAGGGGTGCTTTTAAAGAAATGTTTAA 1769
QY 748 TGGTCTG-TGTTCCAGAAAGCTTTCGACGTACACAGTCCCTTGGGGCAGTAA 806
DB 1770 AGATCGGTTTTTTTAAAGAGGTTTTGTAGTATGGGTTTTGTTGAGGGTAA 1829
QY 807 GCTCTGATTTTCTGTTGCTGCTAGTGAACCCCTACAGAAAGTATGGTGAGCCA 866
DB 1830 GTTTTGAATTTTGTGTTGTTGTTAGTATTTTATTAAGAAATACGGTTAAGTA 1889
QY 867 GG-GGGCGAGCGGCTGCGTCACATGCTGCTGCTTTATCACTTATCATATAAGG 925
DB 1890 GGAAGGCGAGTATTTATTAATATATGTTGTTTATTTATTTATTTATTAAGA 1949
QY 926 AAGGAAAGTATGATTCGATCTGACACTGATGACCTGAGAGGAGAGAAACAAGA 980
DB 1950 AAGGAAAGTATGATTCGATCTGATTTGATGATTTGGGAGAGAGAAATA 2004

RESULT 12
US-10-433-793-188/c
; Sequence 188, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 188
; LENGTH: 2846
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-188

Query Match          11.4%; Score 139.4; DB 7; Length 2846;
Best Local Similarity 56.4%; Pred. No. 2.6e-25;
Matches 457; Conservative 0; Mismatches 316; Indels 38; Gaps 9;

QY 206 CCACAGAGTATGAGCCCGAGAGAAACCTGATACAGTGAAGGAAAGTATGTTTCT 265
DB 1651 CGACAAACGAAACCTCGTCTCAAAAAAAAAAAAAAAAAAATTTATTTCTCTT 1592
```

```

QY 266 TCCTCGACATACCTTCACTCAAAATATCTATTTACTTAAAGACAAATTAACAGCCA 325
DB 1591 TCTACACAAACACTTAATTTCTTAATTTCTCTATTTAATATATTAACCTTAATCAATACA 1532
QY 326 AACCACTTAAAGTTTATTTATGATATTTGTTGTTAAGACAGACATGTGAGTCTG 385
DB 1531 ATATATTTTAAATTAATTAATCAATTAATAATTTTAAATAAATAAATCAAA--AACT 1474
QY 386 AGAAAATCATGTGTGTATCTTGAATTAATA-----TATCAAACTGGGTTAAA 435
DB 1473 CAAAACTAATTTCTTAACGTATTAATAAATAATCTTCTATCAACTATTCGAAA 1414
QY 436 TAAA-----AAAAATGATATCTTAAAAAATAAACAATAACCAAGAACTTTA 488
DB 1413 TAAAAATCTTTATTAATCTTAATTAATTAATCAAAATCAACTTAATTAATCAATAT 1354
QY 489 CTTCCTCTGAAAAAGACA--TTTACAGGGCTGATCTTAGCCTTATATTTACAATA 545
DB 1353 TTTCCCTTAAAAAACAACAGCTTACAAAAAACCAAACTTCTCTTATTTATTAATA 1294
QY 546 AAGAAATTAACCAAGTCCGATATAGCTTAATTTATTCCT-----AAAAGACA 598
DB 1293 ACAAAAATTAACCGAAATCTTAACCTATTTATTAATTTTATTTCTAAAAAATAAACA 1234
QY 599 GAACTTTCATATGCTTTTAAATTAAGTATTAACCTCAGTACTCTGCAAGCTTAGC- 657
DB 1233 AAACTTTCATATTACTTCAACATTTAAATTAATTAACCTCAAAATTTTACCAACTACA 1174
QY 658 CTACAAAGAGAGACAGACAAACAGAGCCCACTACTCTCTAGAAA-----TAATTA 711
DB 1173 GCACAAAAATCAATTTCAACAAAAAATAATCAACTACTCTCTAAAAAATACTTAATTA 1114
QY 712 GGGTGTGCTCTGACATGCCAGGGGTCTTGTGGCTGTCTGTG-TTCCAGAAAGGCTT 770
DB 1113 AAATAATACCTTAATAAATACCAAAATCTTAATCAAAATGATTTTCCAAAAAATTT 1054
QY 771 CTGACGTACACAGTCTTTGGGGAGTAAAGCATATGCTGATTTTCTGTGCTGCTG 830
DB 1053 CTACACATTAATCTTAATTAATAAACAACATTTCTAATTTTCTTAATTAACCTAA 994
QY 831 CTAGTACCCCTTACAGAAATAGTGGTGAACC-AGGGGCGGAGCGGCTGCTGAC 889
DB 993 CTAAATACCCCTTACAAAAAATAACGACTAAACCAAAAAAACAACCAACCTACAC 934
QY 890 ATGTCTGCTGCTCTTATCACTTATCATATATTAAGGAAAGTATGATTTGCGATAC 949
DB 933 ATATCTAATCTACTTATCACTTATCATATTAATAAATAAATAATTAATTAATCGAATAC 874
QY 950 TGACACTGTATACCTCAGGGGAGAAACAAGA 980
DB 873 TAACTATTAATAATCTTAATAAATAAATAAACA 843

RESULT 13
US-10-311-455-866/c
; Sequence 866, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
```

```

; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 866
; LENGTH: 10710
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-866

Query Match
Best Local Similarity 4.5%; Score 55; DB 6; Length 10710;
Matches 148; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 256 TAGTGTGTTCTCTCGACATACCTGACCAATTCATTTGTTACTTAACGACAAAT 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5616 TATTAATTCATTCACGATATCCCTATTAATAAATAAATTCAAAACCAACCT 5557

QY 316 AATCAGGCCAAACCACTTAAGTTTATTTGATAGTATTTGTTAGGACAGACAT 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5556 AATATATTATTCCTTTTCAAAATACATTTGATTTTATATATACCTTACATATTT 5497

QY 376 GTGAGTGTGAGAAAACGATGTTGGTAACTTGATTTAAATATCAAACTGGGTTAAA 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5496 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5437

QY 436 TAAAAAATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5436 AATATCTTCTACATCAATTAACAAAAAATAAATTAATTCCTCCATCCCTC 5377

QY 496 TTGGAAGACACATTTACAAAGGCTGATCTTAGCCCTTATTTTACATTAAGAAATA 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 5376 TAAATTTACAAAATAATTAATTAACCAACAAATTAATTAATTAATTAATTAATTA 5317

QY 556 ACC 558
   |||
DB 5316 AAC 5314

RESULT 14
US-10-311-455-2114/c
; Sequence 2114, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation Status of Cytosine Residues
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 2114
; LENGTH: 15698
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2114

Query Match
Best Local Similarity 4.5%; Score 54.4; DB 6; Length 15698;
Matches 127; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 416 AATATCAAACTGGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 475
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7371 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7312
```

```

QY 476 CAACAAGCTTACTCCCTCGGAAAGCACATTTACAAAGGCTGATCTTACCTTTAT 535
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7311 AACCAATACTACTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 7252

QY 536 ATTTACATTAAGAAATTAACCAAGTCCCGATATAGCTTAATTTATTCCTAAAGA 595
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7251 AATATCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7192

QY 596 ACAGAACTTTCATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 655
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7191 ACTCTACCTAACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 7132

QY 656 GCCTACAA 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7131 CACTACAA 7124

RESULT 15
US-10-602-494-272/c
; Sequence 272, Application US/10602494
; Publication No. US2004026583A1
; GENERAL INFORMATION:
; APPLICANT: Cathy Lofton-Day
; APPLICANT: Andrew Sledziewski
; APPLICANT: Jeff Thomas
; APPLICANT: Robert W. Day
; APPLICANT: Lori Tonnes-Priddy
; APPLICANT: Karen Cardon
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of colorectal cell proliferation
; FILE REFERENCE: 47675-45
; CURRENT APPLICATION NUMBER: US/10/602,494
; CURRENT FILING DATE: 2003-06-23
; NUMBER OF SEQ ID NOS: 385
; SEQ ID NO 272
; LENGTH: 2265
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-602-494-272

Query Match
Best Local Similarity 4.3%; Score 52.6; DB 8; Length 2265;
Matches 115; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 422 AAACGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 222 AAACCAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 163

QY 482 GACTTACTCTCCCTGGAAGACATTTACAAAGGCTGATCTTATATATTTAC 541
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 162 ACATTTACTTAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 103

QY 542 AATTAAGAAATTAACCAAGTCCCGATATAGCTTAATTTATTCCTTAAGAACAGAA 601
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 43

QY 602 ACTTTCATGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 640
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 42 ACTTAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4
```

Search completed: December 12, 2005, 15:40:53
Job time : 1003 secs

THIS PAGE BLANK (USPTO)